WALSH SPACES CONTAINING SMOOTH FUNCTIONS AND QUASI-MONTE CARLO RULES OF ARBITRARY HIGH ORDER

JOSEF DICK[*]

Abstract. We define a Walsh space which contains all functions whose partial mixed derivatives up to order $\delta \geq 1$ exist and have finite variation. In particular, for a suitable choice of parameters, this implies that certain Sobolev spaces are contained in these Walsh spaces. For this Walsh space we then show that quasi-Monte Carlo rules based on digital (t, α, s)-sequences achieve the optimal rate of convergence of the worst-case error for numerical integration. This rate of convergence is also optimal for the subspace of smooth functions. Explicit constructions of digital (t, α, s)-sequences are given hence providing explicit quasi-Monte Carlo rules which achieve the optimal rate of convergence of the integration error for arbitrarily smooth functions.

1. Introduction. Quasi-Monte Carlo rules are quadrature rules which aim to approximate an integral $\int_{[0,1]^s} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$ by the average of the N function values $f(\boldsymbol{x}_n)$ at the quadrature points $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1} \in [0,1]^s$ (and hence are equal weight quadrature rules). The dimension s can be arbitrarily large. The task here is to find ways of how to choose those quadrature points in order to obtain a fast convergence of the approximation to the integral. Explicit constructions of quadrature points in arbitrary high dimensions are until now available for the following two cases:

1. for sufficiently smooth periodic functions arbitrary high convergence can be achieved using Kronecker sequences [18, Theorem 5.3] or a modification of digital nets recently introduced in [2];
2. a convergence of $\mathcal{O}(N^{-1}(\log N)^{s-1})$ can be achieved for functions of bounded variation (in this case the functions are not required to be periodic).

For non-periodic functions no explicit constructions have been established which can fully exploit the smoothness of the integrand. This paper provides a complete solution to this problem.

Among other things we show that an explicit construction of suitable point sets and sequences can be obtained in the following way: let $d \geq 1$ be an integer and let $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots \in [0,1)^{ds}$ be the points of a digital (t, m, ds)-net or digital (t, ds)-sequence over a finite field $\mathbb{F}_q$ in dimension ds (see [20] for the definition of digital nets and sequences and see for example [9, 20, 21, 23, 31] for explicit constructions of suitable digital nets and sequences). Let $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,ds})$ with $x_{n,j} = x_{n,j,1}q^{-1} + x_{n,j,2}q^{-2} + \cdots$ and $x_{n,j,i} \in \{0, \ldots, q-1\}$ (i.e. $x_{n,j,i}$ are the digits in the base q representation of $x_{n,j}$). Then for $n \geq 0$ we define $\boldsymbol{y}_n = (y_{n,1}, \ldots, y_{n,s})$ with

$$y_{n,j} = \sum_{i=1}^{\infty} \sum_{k=1}^{d} x_{n,(j-1)d+k,i} q^{-k-(i-1)d} \quad \text{for } j = 1, \ldots, s.$$

(Note that the addition here is carried out in $\mathbb{R}$ and that the sum over i above is often finite as $x_{n,j,i} = 0$ for i large enough.) We point out here that the quality of the

[*]School of Mathematics and Statistics, University of New South Wales, Sydney 2052, Australia. (`josef.dick@unsw.edu.au`)

point set or sequence is directly related to the t-value of the underlying (t, m, ds)-net or (t, ds)-sequence, see Theorem 4.11 and Theorem 4.12.

Corollary 5.5 now shows that quasi-Monte Carlo rules using the points $\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{N-1}$ (with $N = q^m$ for some $m \geq 1$) achieve the optimal rate of convergence of the integration error of $\mathcal{O}(N^{-\vartheta}(\log N)^{\vartheta s})$ for functions which have partial mixed derivatives up to order ϑ which are square integrable as long as $1 \leq \vartheta \leq d$ (Corollary 5.5 is actually more general). If $\vartheta > d$ no improvement of the convergence rate is obtained compared to functions with smoothness $\vartheta = d$, i.e. we obtain a convergence of $\mathcal{O}(N^{-d}(\log N)^{ds})$. Similar, but less general results for periodic functions compared to those in this paper have been shown in [2] by a different proof method. (The construction above is an example of a construction method which can be used. In Section 4 we outline the general algebraical properties required for the construction of suitable point sets.)

The quasi-Monte Carlo algorithm based on digital nets and sequences proposed here has also some further useful properties. For example our results also hold if one randomizes the point set by, say, a random digital shift (see for example [4, 5, 17]). (This follows easily because the worst-case error (see Section 5) is invariant with respect to digital shifts in the Walsh space and hence we obtain the same upper bounds for randomized digital nets and sequences.) In summary the quadrature rules have the following properties:

- The quadrature rules introduced in this paper are equal weight quadrature rules which achieve the optimal rate of convergence up to some $\log N$ factors and the result holds for deterministic and randomly digitally shifted quadrature rules.
- The construction of the underlying point set is explicit and suitable point sets are available in arbitrary high dimensions and arbitrary high number of points.
- The quadrature rules automatically adjust themselves to the optimal rate of convergence $\mathcal{O}(N^{-\vartheta}(\log N)^{s\vartheta})$ as long as $1 \leq \vartheta \leq d$.
- The underlying point set is extensible in the dimension as well as in the number of points, i.e., one can always add some coordinates or points to an existing point set such that the quality of the point set is preserved.

In the following we lay out some of the underlying principles used in this work which stem from the behaviour of the Walsh coefficients of smooth functions. Walsh functions are piecewise constant wavelets which form an orthonormal set of $\mathcal{L}_2([0, 1]^s)$. In their simplest form, for a non-negative integer k with base 2 representation $k = \kappa_0 + \cdots + \kappa_{m-1} 2^{m-1}$ and an $x \in [0, 1)$ with base 2 representation $x = x_1 2^{-1} + x_2 2^{-2} + \cdots$, the k-th Walsh function in base 2 is given by

$$\mathrm{wal}_k(x) = (-1)^{\kappa_0 x_1 + \cdots + \kappa_{m-1} x_m}.$$

(Later on we will use the more general definition of Walsh functions over groups.)

The behaviour of the Fourier coefficients of smooth periodic functions is well known, i.e. the smoother the function the faster the Fourier coefficients go to zero (see for example [34]). An analogous result for Walsh functions has, to the best of the authors knowledge, not been known until now (see Fine [10] who, for example, shows that the only absolute continuous functions whose k-th Walsh coefficients decay faster than $1/k$ are constant functions). This will be established here and subsequently be exploited to obtain quasi-Monte Carlo rules with arbitrary high order of convergence.

To give a glimpse of how the Walsh coefficients of smooth functions behave, consider for example the Walsh series for $1/2 - x$:

$$1/2 - x = \sum_{k=0}^{\infty} c_k \text{wal}_k(x) = \sum_{a=0}^{\infty} 2^{-a-2} \text{wal}_{2^a}(x).$$

Although the function is infinitely smooth, in general the decay of the Walsh coefficient is only of order $1/k$. But note that most of the Walsh coefficients are actually 0. For example when we consider $(1/2 - x)^2$, then typically we would have that the Walsh coefficient of $k = 2^a$ is of order 2^{-a}, the Walsh coefficient of $k = 2^{a_1} + 2^{a_2}$ $(a_1 > a_2)$ is of order $2^{-a_1-a_2}$ and for $k = 2^{a_1} + 2^{a_2} + \cdots + 2^{a_v}$ with $a_1 > \cdots > a_v$ and $v > 2$ the k-th Walsh coefficient would be 0. By considering $(1/2 - x)^3, (1/2 - x)^4, \ldots$, or more generally polynomials, one can now realize that the speed of convergence of the Walsh coefficients depends on how many non-zero digits k has. This is the basic feature which we will relate to the speed of convergence of the Walsh coefficients for smooth functions.

Subsequently we will explicitly state and use the behaviour of the Walsh coefficients of smooth functions. In general the Walsh functions depend on the base q digit expansion of the wavenumber k and also of the point x where the Walsh function is to be evaluated. Hence, maybe not surprisingly, the value of the k-th Walsh coefficients of smooth functions also depend on the q-adic expansion of k. We show that the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ introduced in Section 3 contains all functions whose partial mixed derivatives up to order $\delta < \vartheta$ exist and have finite variation, where ϑ is a parameter restricting the behaviour of the Walsh coefficients of the function space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$. (We use a similar, though much more general, technique as Fine [10] used for showing that the Walsh coefficients of a differentiable function cannot decay faster than $1/k$.)

The concept of digital (t, α, β, m, s)-nets and digital (t, α, β, s)-sequences (see Section 4 and also [2] for a similar concept) is now designed to yield point sets which work well for the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$, just in the same way as the digital nets and sequences from [9, 19, 20, 23, 31] are designed to work well for the spaces for example considered in [4, 11] (or as lattice rules are designed to work well for periodic Korobov spaces). Here the power of the result that the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ contains smooth functions comes into play: it follows that we can fully exploit the smoothness of an integrand using digital $(t, \alpha, \alpha, m, s)$-nets or digital (t, α, α, s)-sequences. As the construction of the points $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots$ introduced at the beginning yields explicit examples of digital $(t, \alpha, \alpha, m, s)$-nets or digital (t, α, α, s)-sequences as shown in Section 4 we therefore obtain explicit constructions of quasi-Monte Carlo rules which can achieve the optimal order of convergence for arbitrary smooth functions.

In the next section we introduce Walsh functions over groups and state some of their essential properties.

2. Walsh functions over groups. In this section we give the definition of Walsh functions over groups and present some essential properties. Walsh functions in base 2 were first introduced by Walsh [33], though a similar but non-complete set of functions has already been studied by Rademacher [27]. Further important results were obtained in [10]. We follow [26] in our presentation.

2.1. Definition of Walsh functions over groups. An essential tool for the investigation of digital nets are Walsh functions. A very general definition, corresponding to the most general construction of digital nets over finite rings, was given

in [15]. There, Walsh functions over a finite abelian group G, using some bijection φ, were defined. Here we restrict ourselves to the additive groups of the finite fields $\mathbb{F}_{p^r}$, p prime and $r \geq 1$. We restate the definitions for this special case here for the sake of convenience. In the following let $\mathbb{N}$ denote the set of positive integers and $\mathbb{N}_0$ the set of non-negative integers.

DEFINITION 2.1 (Walsh functions). Let $q = p^r$, p prime, $r \in \mathbb{N}$ and let $\mathbb{F}_q$ be the finite field with q elements. Let $\mathbb{Z}_q = \{0, 1, \dots, q-1\} \subset \mathbb{Z}$ and let $\varphi : \mathbb{Z}_q \longrightarrow \mathbb{F}_q$ be a bijection such that $\varphi(0) = 0$, the neutral element of addition in $\mathbb{F}_q$. Moreover denote by ψ the canonical isomorphism (described below) of additive groups $\psi : \mathbb{F}_q \longrightarrow \mathbb{Z}_p^r$ and define $\eta := \psi \circ \varphi$. For $1 \leq i \leq r$ denote by π_i the projection $\pi_i : \mathbb{Z}_p^r \longrightarrow \mathbb{Z}_p$, $\pi_i(x_1, \dots, x_r) = x_i$.

$$
\begin{array}{ccc}
\mathbb{Z}_q & \xrightarrow{\;\varphi\;} & \mathbb{F}_q \\
& \eta \searrow & \downarrow \psi \\
& & \mathbb{Z}_p^r \xrightarrow{\;\pi_i\;} \mathbb{Z}_p
\end{array}
$$

Let now $k \in \mathbb{N}_0$ with base q representation $k = \kappa_0 + \kappa_1 q + \cdots + \kappa_{m-1} q^{m-1}$ where $\kappa_l \in \mathbb{Z}_q$ and let $x \in [0,1)$ with base q representation $x = x_1/q + x_2/q^2 + \cdots$ (unique in the sense that infinitely many x_l must be different from $q-1$). Then the k-th Walsh function over the additive group of the finite field $\mathbb{F}_q$ with respect to the bijection φ is defined by

$$
{}_{\mathbb{F}_q,\varphi}\mathrm{wal}_k(x) = \exp\left(\frac{2\pi \mathtt{i}}{p} \sum_{l=0}^{m-1} \sum_{i=1}^{r} (\pi_i \circ \eta)(\kappa_l)(\pi_i \circ \eta)(x_{l+1}) \right).
$$

For convenience we will in the rest of the paper omit the subscript and simply write wal_k if there is no ambiguity.

Multivariate Walsh functions are defined by multiplication of the univariate components, i.e., for $s > 1$, $\boldsymbol{x} = (x_1, \dots, x_s) \in [0,1)^s$ and $\boldsymbol{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s$, we set

$$
\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) = \prod_{j=1}^{s} \mathrm{wal}_{k_j}(x_j).
$$

We now briefly describe the canonical isomorphism. Let $\mathbb{F}_q = \mathbb{Z}_p[\theta]$, such that $\{1, \theta, \dots, \theta^{r-1}\}$ is a basis of $\mathbb{F}_q$ over $\mathbb{Z}_p$ as a vector space. Then the isomorphism ψ between $\mathbb{F}_q$ and $\mathbb{Z}_p^r$ shall be given by

$$
\psi(x) = (x_1, \dots, x_r)^\top, \text{ for } x = \sum_{i=1}^{r} x_i \theta^{i-1}, x_i \in \mathbb{Z}_p.
$$

For more information on the Walsh functions defined above see [26].

We summarize some important properties of Walsh functions over the additive group of a finite field which will be used throughout the paper. The proofs of the subsequent results can be found e.g. in [16, 25] (see also [1]). In the following we call $x \in [0,1)$ a q-adic rational if x can be represented by a finite base q expansion.

PROPOSITION 2.2. *Let p, q, $\mathbb{F}_q$ and φ be as in Definition 2.1. For x, y with q-adic representations $x = \sum_{i=w}^{\infty} x_i q^{-i}$ and $y = \sum_{i=w}^{\infty} y_i q^{-i}$, $w \in \mathbb{Z}$ (taking w negative, hence the following operations are also defined for integers), define $x \oplus_\varphi y :=$*

$\sum_{i=w}^{\infty} z_i q^{-i}$ where $z_i := \varphi^{-1}(\varphi(x_i) + \varphi(y_i))$ and $\ominus_{\varphi} x := \sum_{i=w}^{\infty} v_i q^{-i}$ where $v_i := \varphi^{-1}(-\varphi(x_i))$. Further we set $x \ominus_{\varphi} y := x \oplus_{\varphi} (\ominus_{\varphi} y)$. For vectors $\boldsymbol{x}, \boldsymbol{y}$ we define the operations component-wise. Then we have:

1. For all $k, l \in \mathbb{N}_0$ and all $x, y \in [0, 1)$, with the restriction that if x, y are not q-adic rationals then $x \oplus_{\varphi} y$ is not allowed to be a q-adic rational, we have

$$\mathrm{wal}_k(x) \cdot \mathrm{wal}_l(x) = \mathrm{wal}_{k \oplus_{\varphi} l}(x), \quad \mathrm{wal}_k(x) \cdot \mathrm{wal}_k(y) = \mathrm{wal}_k(x \oplus_{\varphi} y)$$

and, with the restriction that if x, y are not q-adic rationals then $x \ominus_{\varphi} y$ is not allowed to be a q-adic rational,

$$\mathrm{wal}_k(x) \cdot \overline{\mathrm{wal}_l(x)} = \mathrm{wal}_{k \ominus_{\varphi} l}(x), \quad \mathrm{wal}_k(x) \cdot \overline{\mathrm{wal}_k(y)} = \mathrm{wal}_k(x \ominus_{\varphi} y).$$

2. We have

$$\sum_{k=0}^{q-1} \mathrm{wal}_l(k/q) = \begin{cases} 0 & if\ l \neq 0, \\ q & if\ l = 0. \end{cases}$$

3. We have

$$\int_0^1 \mathrm{wal}_0(x)\, \mathrm{d}x = 1 \quad and \quad \int_0^1 \mathrm{wal}_k(x)\, \mathrm{d}x = 0\ if\ k > 0.$$

4. For all $\boldsymbol{k}, \boldsymbol{l} \in \mathbb{N}_0^s$ we have the following orthogonality properties:

$$\int_{[0,1)^s} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) \overline{\mathrm{wal}_{\boldsymbol{l}}(\boldsymbol{x})}\, \mathrm{d}\boldsymbol{x} = \begin{cases} 1 & if\ \boldsymbol{k} = \boldsymbol{l}, \\ 0 & otherwise. \end{cases}$$

5. For any $f \in \mathcal{L}_2([0,1)^s)$ and any $\boldsymbol{\sigma} \in [0, 1)^s$ we have

$$\int_{[0,1)^s} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \int_{[0,1)^s} f(\boldsymbol{x} \oplus_{\varphi} \boldsymbol{\sigma})\, \mathrm{d}\boldsymbol{x}.$$

6. For any integer $s \geq 1$ the system $\{\mathrm{wal}_{\boldsymbol{k}} : \boldsymbol{k} \in \mathbb{N}_0^s\}$ is a complete orthonormal system in $\mathcal{L}_2([0, 1)^s)$.

REMARK 2.3. The restrictions in item 1. was added to exclude cases like: $x = (0.010101\ldots)_2$, $y = (0.0010101\ldots)_2$ and $x \oplus y = (0.1)_2$, for which the result is of course not true. On the other hand, the result holds for $x \oplus y = (0.0111111\ldots)_2$.

Throughout the paper we will use a fixed bijection φ and a fixed finite field $\mathbb{F}_q$ is used for Walsh functions and $\oplus_{\varphi}$ and $\ominus_{\varphi}$. Hence we will often write $\oplus$ and $\ominus$ instead of $\oplus_{\varphi}$ and $\ominus_{\varphi}$.

In the following section we will deal with Walsh series and Walsh coefficients, which we briefly describe in the following: functions $f \in \mathcal{L}_2([0, 1)^s)$ have an associated Walsh series

$$f(\boldsymbol{x}) \sim \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} \hat{f}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}),$$

where the Walsh coefficients $\hat{f}(\boldsymbol{k})$ are given by

$$\hat{f}(\boldsymbol{k}) = \int_{[0,1)^s} f(\boldsymbol{x}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

For smooth functions the Walsh series converges to the function, which is shown in Section 3.3.

3. Walsh spaces containing smooth functions. In the following we investigate how the Walsh coefficients of smooth functions decay and subsequently we use this to define function classes based on Walsh functions which contain smooth functions. But first we introduce a suitable variation.

3.1. A generalized weighted Hardy and Krause variation. In the following we generalize the Hardy and Krause variation which suits our purposes later on.

3.1.1. Hölder condition. A function $f : [0, 1) \to \mathbb{R}$ satisfies a Hölder condition with coefficient $0 < \lambda \le 1$ if there is a constant $C_f > 0$ such that

$$|f(x) - f(y)| \le C_f |x - y|^\lambda \quad \text{for all } x, y \in [0, 1).$$

The right hand side of the above inequality forms a metric on $[0, 1)$. When one considers the higher dimensional domain $[0, 1)^s$ then $|x - y|$ is changed to some other metric on $[0, 1)^s$. Here we consider tensor product spaces and we generalize the Hölder condition to higher dimensions in a way which is suitable for tensor product spaces in our context. Consider for example the function $f(\boldsymbol{x}) = \prod_{j=1}^{s} f_j(x_j)$, where $\boldsymbol{x} = (x_1, \ldots, x_s)$ and each $f_j : [0, 1) \to \mathbb{R}$ satisfies a Hölder condition with coefficient $0 < \lambda \le 1$. Then it follows that for all $\emptyset \ne u \subseteq \mathcal{S} := \{1, \ldots, s\}$ we have

$$\prod_{j \in u} |f_j(x_j) - f_j(y_j)| \le \prod_{j \in u} C_{f_j} \prod_{j \in u} |x_j - y_j|^\lambda \tag{3.1}$$

for all $x_j, y_j \in [0, 1)$ with $j \in u$. But here $\prod_{j=1}^{s} |x_j - y_j|$ is not a metric on $[0, 1)^s$.

Note that we have

$$\prod_{j \in u} |f_j(x_j) - f_j(y_j)| = \left| \sum_{v \subseteq u} (-1)^{|v| - |u|} \prod_{j \in v} f_j(x_j) \prod_{j \in u \setminus v} f_j(y_j) \right|, \tag{3.2}$$

which can be described in words in the following way: for given $\emptyset \ne u \subseteq \mathcal{S}$ let $x_j, y_j \in [0, 1)$ with $x_j \ne y_j$ for all $j \in u$; consider the box J with vertices $\{(a_j)_{j \in u} : a_j = x_j \text{ or } a_j = y_j \text{ for } j \in u\}$. Then (3.2) is the alternating sum of the function $\prod_{j \in u} f_j$ at the vertices of J where adjacent vertices have opposite signs. This sum can also be defined for functions on $[0, 1)^s$ which are not of product form.

Indeed, let for a subinterval $J = \prod_{j=1}^{s} [x_j, y_j)$ with $0 \le x_j < y_j \le 1$ and a function $f : [0, 1)^s \to \mathbb{R}$ the function $\Delta(f, J)$ denote the alternating sum of f at the vertices of J where adjacent vertices have opposite signs. (Hence for $f = \prod_{j=1}^{s} f_j$ we have $\Delta(f, J) = \prod_{j=1}^{s} (f_j(x_j) - f_j(y_j))$.)

3.1.2. Generalized Vitali variation. Let $\mathfrak{p} \ge 1$. Then we define the generalized variation in the sense of Vitali with coefficient $0 < \lambda \le 1$ by

$$V_{\lambda, \mathfrak{p}}^{(s)}(f) = \sup_{\mathcal{P}} \left(\sum_{J \in \mathcal{P}} \text{Vol}(J) \left| \frac{\Delta(f, J)}{\text{Vol}(J)^\lambda} \right|^{\mathfrak{p}} \right)^{1/\mathfrak{p}}, \tag{3.3}$$

where the supremum is extended over all partitions $\mathcal{P}$ of $[0, 1]^s$ into subintervals and $\text{Vol}(J)$ denotes the volume of the subinterval J.

Note that for $\lambda = 1$ and $\mathfrak{p} = 1$ one obtains the usual definition of the Vitali variation, see for example [20]. If we take $\mathfrak{p} = \infty$, then we obtain a condition of the form (3.1) where $u = \mathcal{S}$ and where we can take the constant $\prod_{j=1}^{s} C_{f_j} = V_{\lambda, \infty}^{(s)}(f)$.

For $s = 1$ and $\mathfrak{p} = \infty$ we obtain a Hölder condition with coefficient $0 < \lambda \leq 1$. In this sense we can view (3.3) as a fractional Vitali variation of order λ.

For $\lambda = 1$ and if the partial derivatives of f are continuous on $[0, 1]^s$ we also have the formula

$$V_{1,\mathfrak{p}}^{(s)}(f) = \left(\int_{[0,1]^s} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s} \right|^{\mathfrak{p}} \mathrm{d}\boldsymbol{x} \right)^{1/\mathfrak{p}}, \tag{3.4}$$

for all $\mathfrak{p} \geq 1$. Indeed we have

$$|\Delta(f, J)| = \left| \int_J \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right| = \mathrm{Vol}(J) \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(\boldsymbol{\zeta}_J) \right|$$

for some $\boldsymbol{\zeta}_J \in \overline{J}$, which follows by applying the mean value theorem to the inequality

$$\min_{\boldsymbol{x} \in \overline{J}} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(\boldsymbol{x}) \right| \leq \mathrm{Vol}(J)^{-1} \left| \int_J \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right| \leq \max_{\boldsymbol{x} \in \overline{J}} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(\boldsymbol{x}) \right|.$$

Therefore we have

$$\sum_{J \in \mathcal{P}} \mathrm{Vol}(J) \left| \frac{\Delta(f, J)}{\mathrm{Vol}(J)} \right|^{\mathfrak{p}} = \sum_{J \in \mathcal{P}} \mathrm{Vol}(J) \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(\boldsymbol{\zeta}_J) \right|^{\mathfrak{p}},$$

which is just a Riemann sum for the integral $\int_{[0,1]^s} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s} \right|^{\mathfrak{p}} \mathrm{d}\boldsymbol{x}$ and thus the equality follows.

Using Hölder's inequality and the fact that $\left(\sum_{J \in \mathcal{P}} (\mathrm{Vol}(J)^{1-1/\mathfrak{p}})^{\mathfrak{p}/(\mathfrak{p}-1)} \right)^{1-1/\mathfrak{p}} = \left(\sum_{J \in \mathcal{P}} \mathrm{Vol}(J) \right)^{1-1/\mathfrak{p}} = 1$ it follows that

$$V_{\lambda,1}^{(s)}(f) \leq V_{\lambda,\mathfrak{p}}^{(s)}(f) \quad \text{for all } \mathfrak{p} \geq 1.$$

3.1.3. Generalized Hardy and Krause variation. Until now we did not take projections to lower dimensional faces into account (in (3.1) we did take projections into account as we considered all $\emptyset \neq u \subseteq \mathcal{S}$).

For $\emptyset \neq u \subseteq \mathcal{S}$, let $V_{\lambda,\mathfrak{p}}^{(|u|)}(f_u; u)$ be the generalized Vitali variation with coefficient $0 < \lambda \leq 1$ of the $|u|$-dimensional function $f_u(\boldsymbol{x}_u) = \int_{[0,1)^{s-|u|}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{\mathcal{S} \setminus u}$. For $u = \emptyset$ we have $f_\emptyset = \int_{[0,1)^s} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_\mathcal{S}$ and we define $V_{\lambda,\mathfrak{p}}^{(|\emptyset|)}(f_\emptyset; \emptyset) = |f_\emptyset|$. Let $\mathfrak{q} \geq 1$, then

$$V_{\lambda,\mathfrak{p},\mathfrak{q}}(f) = \left(\sum_{u \subseteq \mathcal{S}} \left(V_{\lambda,\mathfrak{p}}^{(|u|)}(f_u; u) \right)^{\mathfrak{q}} \right)^{1/\mathfrak{q}} \tag{3.5}$$

is called the generalized Hardy and Krause variation of f on $[0, 1]^s$.

For $\lambda = \mathfrak{p} = \mathfrak{q} = 1$ one obtains an unanchored version of the usual definition of the Hardy and Krause variation, see [20]. A function f for which $V_{\lambda,\mathfrak{p},\mathfrak{q}}(f) < \infty$ is said to be of finite variation with coefficient λ. (We remark that in some cases it might be appropriate to leave out the term corresponding to $u = \emptyset$ in (3.5), but here this term will be needed later on and hence we include it already in the definition of the variation.)

3.1.4. Generalized weighted Hardy and Krause variation. As first suggested in [30] (see also [6]) different coordinates might have different importance, hence we can also define a weighted variation. In the spirit of the weighted Sobolev spaces in [6], let $\boldsymbol{\gamma} = (\gamma_u)_{u \subset \mathbb{N}}$ be an indexed set of non-negative real numbers. Then we define the weighted variation $V_{\lambda,\mathfrak{p},\mathfrak{q},\boldsymbol{\gamma}}(f)$ of f with coefficient $0 < \lambda \leq 1$ by

$$V_{\lambda,\mathfrak{p},\mathfrak{q},\boldsymbol{\gamma}}(f) = \left(\sum_{u \subseteq \mathcal{S}} \gamma_u^{-1} \left(V_{\lambda,\mathfrak{p}}^{(|u|)}(f_u; u) \right)^{\mathfrak{q}} \right)^{1/\mathfrak{q}}.$$

Note that for $\lambda = 1$ and $\mathfrak{p} = \mathfrak{q} = 2$ the weighted variation $V_{\lambda,\mathfrak{p},\mathfrak{q},\boldsymbol{\gamma}}(f)$ coincides with the norm in a weighted unanchored Sobolev space for any function in this Sobolev space, i.e, we have the identity $V_{1,2,2,\boldsymbol{\gamma}}(f) = \|f\|_{\mathrm{sob}}$, where

$$\|f\|_{\mathrm{sob}} = \left(\sum_{u \subseteq \mathcal{S}} \gamma_u^{-1} \int_{[0,1)^{|u|}} \left| \int_{[0,1)^{s-|u|}} \frac{\partial^{|u|} f(\boldsymbol{x})}{\partial \boldsymbol{x}_u} \, \mathrm{d}\boldsymbol{x}_{\mathcal{S} \setminus u} \right|^2 \mathrm{d}\boldsymbol{x}_u \right)^{1/2}$$

denotes the norm in the weighted Sobolev space (see [6] for more information on this Sobolev space).

3.2. The decay of the Walsh coefficients of smooth functions. We are now ready to show how the Walsh coefficients of smooth functions decay. This behaviour is essentially captured in Definition 3.5 below. But before we get there we need several lemmas to prove the result. The following lemma is needed to show how the Walsh coefficients of functions with bounded variation decay. A simpler version of it was shown in [25, Lemma 4].

LEMMA 3.1. *Let* $f \in \mathcal{L}_1([0,1)^s)$ *and let* $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}^s$ *with* $k_j = \kappa_j q^{a_j - 1} + k_j'$ *where* $a_j \in \mathbb{N}$, $\kappa_j \in \{1, \ldots, q-1\}$, $0 \leq k_j' < q^{a_j - 1}$ *and let* $0 \leq c_j < q^{a_j - 1}$ *for* $j = 1, \ldots, s$. *Then*

$$\left| \int_{\prod_{j=1}^s [c_j q^{-a_j + 1}, (c_j + 1)q^{-a_j + 1})} f(\boldsymbol{x}) \, \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x} \right| \leq q^{-\sum_{j=1}^s (a_j - 1)} \sup_J |\Delta(f, J)|,$$

where the supremum is taken over all boxes of the form

$$J = \prod_{j=1}^s [d_j, e_j) \subseteq \prod_{j=1}^s [c_j q^{-a_j + 1}, (c_j + 1)q^{-a_j + 1})$$

with $q^{a_j} |e_j - d_j| \in \{1, \ldots, q-1\}$.

Proof. We have $\overline{\mathrm{wal}_{k_j}} = \overline{\mathrm{wal}_{\kappa_j q^{a_j - 1}}} \, \overline{\mathrm{wal}_{k_j'}}$ and the function $\overline{\mathrm{wal}_{k_j'}}$ is constant on each subinterval $[c_j q^{-a_j + 1}, (c_j + 1)q^{-a_j + 1})$. Hence we have

$$\left| \int_{\prod_{j=1}^s [c_j q^{-a_j + 1}, (c_j + 1)q^{-a_j + 1})} f(\boldsymbol{x}) \, \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x} \right|$$

$$= \left| \int_{\prod_{j=1}^s [c_j q^{-a_j + 1}, (c_j + 1)q^{-a_j + 1})} f(\boldsymbol{x}) \prod_{j=1}^s \overline{\mathrm{wal}_{\kappa_j q^{a_j - 1}}(x_j)} \, \mathrm{d}\boldsymbol{x} \right|.$$

Note that the function $\overline{\mathrm{wal}_{\kappa_j q^{a_j - 1}}}$ is constant on each of the subintervals $[r_j q^{-a_j}, (r_j + 1)q^{-a_j})$ for $r_j = 0, \ldots, q^{a_j} - 1$ for $j = 1, \ldots, s$. Without loss of generality we may

assume that $c_j = 0$, for all other c_j the result follows by the same arguments. Thus we have

$$\int_{\prod_{j=1}^s [0, q^{-a_j+1})} f(\boldsymbol{x}) \prod_{j=1}^s \overline{\mathrm{wal}_{\kappa_j q^{a_j-1}}(x_j)} \, \mathrm{d}\boldsymbol{x}$$

$$= \sum_{r_1,\ldots,r_s=0}^{q-1} \prod_{j=1}^s \overline{\mathrm{wal}_{\kappa_j}(r_j/q)} \int_{\prod_{j=1}^s [r_j q^{-a_j}, (r_j+1)q^{-a_j})} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

Let now $a_{(r_1,\ldots,r_s)} = \int_{\prod_{j=1}^s [r_j q^{-a_j}, (r_j+1)q^{-a_j})} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$ and for a given $0 \le r_1,\ldots,r_s < q$ let

$$A(r_1,\ldots,r_s) = q^{-s} \sum_{t_1,\ldots,t_s=0}^{q-1} \sum_{\emptyset \neq u \subseteq \{1,\ldots,s\}} (-1)^{|u|} a_{(\boldsymbol{t}_u, \boldsymbol{r}_{\{1,\ldots,s\}\setminus u})}, \qquad (3.6)$$

where $(\boldsymbol{t}_u, \boldsymbol{r}_{\{1,\ldots,s\}\setminus u})$ denotes the vector obtained by setting the j-th coordinate to t_j if $j \in u$ and r_j if $j \notin u$. Further let

$$B(r_1,\ldots,r_s) = q^{-s} \sum_{t_1,\ldots,t_s=0}^{q-1} \sum_{u \subseteq \{1,\ldots,s\}} (-1)^{|u|} a_{(\boldsymbol{t}_u, \boldsymbol{r}_{\{1,\ldots,s\}\setminus u})}.$$

Then we have

$$\sum_{r_1,\ldots,r_s=0}^{q-1} \prod_{j=1}^s \overline{\mathrm{wal}_{\kappa_j}(r_j/q)} a_{(r_1,\ldots,r_s)}$$

$$= - \sum_{r_1,\ldots,r_s=0}^{q-1} \prod_{j=1}^s \overline{\mathrm{wal}_{\kappa_j}(r_j/q)} A(r_1,\ldots,r_s)$$

$$+ \sum_{r_1,\ldots,r_s=0}^{q-1} \prod_{j=1}^s \overline{\mathrm{wal}_{\kappa_j}(r_j/q)} (a_{(r_1,\ldots,r_s)} + A(r_1,\ldots,r_s)).$$

Since $\sum_{r=0}^{q-1} \overline{\mathrm{wal}_\kappa(r/q)} = 0$ and $A(r_1,\ldots,r_s)$ is a sum where each summand does not depend on at least one r_j, i.e. the case $u = \emptyset$ is excluded in (3.6), it follows that the first sum on the right hand side above is zero. Further we have $a_{(r_1,\ldots,r_s)} + A(r_1,\ldots,r_s) = B(r_1,\ldots,r_s)$ and thus

$$\left| \sum_{r_1,\ldots,r_s=0}^{q-1} \prod_{j=1}^s \overline{\mathrm{wal}_{\kappa_j}(r_j/q)} a_{(r_1,\ldots,r_s)} \right| \le \sum_{r_1,\ldots,r_s=0}^{q-1} |B(r_1,\ldots,r_s)|.$$

We have

$$|B(r_1,\ldots,r_s)| \le \max_{\boldsymbol{t} \in \{0,\ldots,q-1\}^s} \left| \sum_{u \subseteq \{1,\ldots,s\}} (-1)^{|u|} a_{(\boldsymbol{t}_u, \boldsymbol{r}_{\{1,\ldots,s\}\setminus u})} \right|.$$

Therefore we have

$$\left| \int_{\prod_{j=1}^{s}[0,q^{-a_j+1})} f(\boldsymbol{x}) \prod_{j=1}^{s} \overline{\mathrm{wal}_{\kappa_j q^{a_j-1}}(x_j)} \, \mathrm{d}\boldsymbol{x} \right|$$

$$\leq \sum_{r_1,\ldots,r_s=0}^{q-1} \max_{\boldsymbol{t}\in\{0,\ldots,q-1\}^s} \left| \sum_{u\subseteq\{1,\ldots,s\}} (-1)^{|u|} a_{(\boldsymbol{t}_u,\boldsymbol{r}_{\{1,\ldots,s\}\setminus u})} \right|$$

$$\leq q^s \max_{\boldsymbol{r},\boldsymbol{t}\in\{0,\ldots,q-1\}^s} \left| \sum_{u\subseteq\{1,\ldots,s\}} (-1)^{|u|} a_{(\boldsymbol{t}_u,\boldsymbol{r}_{\{1,\ldots,s\}\setminus u})} \right|.$$

Note that if in the above maximum there is a j such that $r_j = t_j$ then it follows that

$$\left| \sum_{u\subseteq\{1,\ldots,s\}} (-1)^{|u|} a_{(\boldsymbol{t}_u,\boldsymbol{r}_{\{1,\ldots,s\}\setminus u})} \right| = 0.$$

Hence we may in the following assume without loss of generality that the maximum in the last line of the inequality above is taken on for $\boldsymbol{r} = (r_1,\ldots,r_s)$ and $\boldsymbol{t} = (t_1,\ldots,t_s)$ which satisfy $r_j \neq t_j$ for $j = 1,\ldots,s$.

We have

$$\left| \sum_{u\subseteq\{1,\ldots,s\}} (-1)^{|u|} a_{(\boldsymbol{t}_u,\boldsymbol{r}_{\{1,\ldots,s\}\setminus u})} \right|$$

$$= \left| \int_{\prod_{j=1}^{s}[r_j q^{-a_j},(r_j+1)q^{-a_j})} \sum_{u\subseteq\{1,\ldots,s\}} (-1)^{|u|} f(\boldsymbol{x} + \boldsymbol{y}_u) \, \mathrm{d}\boldsymbol{x} \right|,$$

where $\boldsymbol{y}_u = (y_1,\ldots,y_s)$ with $y_j = 0$ for $j \notin u$ and $y_j = (t_j - r_j)q^{-a_j}$ for $j \in u$. We can write

$$\sum_{u\subseteq\{1,\ldots,s\}} (-1)^{|u|} f(\boldsymbol{x} + \boldsymbol{y}_u) = \Delta(f, J_{\boldsymbol{x}}),$$

where $J_{\boldsymbol{x}} = \prod_{j=1}^{s}[\min(x_j, x_j + (t_j - r_j)q^{-a_j}), \max(x_j, x_j + (t_j - r_j)q^{-a_j}))$. Therefore it follows that

$$\left| \sum_{u\subseteq\{1,\ldots,s\}} (-1)^{|u|} a_{(\boldsymbol{t}_u,\boldsymbol{r}_{\{1,\ldots,s\}\setminus u})} \right| \leq q^{-\sum_{j=1}^{s} a_j} \sup_{\boldsymbol{x}\in\prod_{j=1}^{s}[r_j q^{-a_j},(r_j+1)q^{-a_j})} |\Delta(f, J_{\boldsymbol{x}})|.$$

The result follows. □

In the following lemma we now obtain a bound on the Walsh coefficients for functions of bounded variation. It is a generalization of [25, Proposition 6].

LEMMA 3.2. *Let $0 < \lambda \leq 1$ and let $f \in \mathcal{L}_2([0,1)^s)$ satisfy $V_{\lambda,1,1,\boldsymbol{\gamma}}(f) < \infty$. Then for any $\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}$ the $\boldsymbol{k}$-th Walsh coefficient of f satisfies*

$$|\hat{f}(\boldsymbol{k})| \leq q^{|u|-\lambda\sum_{j\in u}(a_j-1)} V_{\lambda,1}^{(|u|)}(f_u; u),$$

where $\boldsymbol{k} = (k_1,\ldots,k_s)$, $u = \{1 \leq j \leq s : k_j \neq 0\}$ and for $j \in u$ we have $k_j = \kappa_j q^{a_j-1} + k_j'$, where $\kappa_j \in \{1,\ldots,q-1\}$, $a_j \in \mathbb{N}$ and $0 \leq k_j' < q^{a_j-1}$.

Proof. Let $f \in \mathcal{L}_2([0,1)^s)$ with $\boldsymbol{k}$-th Walsh coefficient $\hat{f}(\boldsymbol{k})$. First note that it suffices to show the result for $\boldsymbol{k} \in \mathbb{N}^s$, as otherwise we only need to replace the function f with the function $f_u(\boldsymbol{x}_u) = \int_{[0,1)^{s-|u|}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{\mathcal{S}\setminus u}$. Hence let now $\boldsymbol{k} \in \mathbb{N}^s$ be given and let $\boldsymbol{k}' = (k_1', \ldots, k_s')$. Then we have

$$\left| \int_{[0,1)^s} f(\boldsymbol{x}) \, \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x} \right|$$

$$\leq \sum_{\substack{0 \leq c_j < q^{a_j} - 1 \\ 1 \leq j \leq s}} \left| \int_{c_1 q^{-a_1+1}}^{(c_1+1)q^{-a_1+1}} \cdots \int_{c_s q^{-a_s+1}}^{(c_s+1)q^{-a_s+1}} f(\boldsymbol{x}) \, \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x} \right|.$$

Now we use Lemma 3.1 and thereby obtain that the above sum is bounded by

$$\sum_{0 \leq c_1 < q^{a_1}-1} \cdots \sum_{0 \leq c_s < q^{a_s}-1} q^{-\sum_{j=1}^{s}(a_j-1)} \sup_J |\Delta(f, J)|,$$

where the supremum is taken over all boxes $J = \prod_{j=1}^s [d_j, e_j) \subseteq \prod_{j=1}^s [c_j q^{-a_j+1}, (c_j + 1)q^{-a_j+1})$ with $q^{a_j}|e_j - d_j| \in \{1, \ldots, q-1\}$. Now we have

$$q^{-\sum_{j=1}^{s}(a_j-1)} \sup_J |\Delta(f, J)| \leq \sup_{\mathcal{P}_{\boldsymbol{c}}} \sum_{I \in \mathcal{P}_{\boldsymbol{c}}} \mathrm{Vol}(I)^{1-\lambda} |\Delta(f, I)| \frac{q^{-\sum_{j=1}^{s}(a_j-1)}}{\mathrm{Vol}(I)^{1-\lambda}},$$

where the supremum on the right hand side is taken over all partitions $\mathcal{P}_{\boldsymbol{c}}$ of the cube $\prod_{j=1}^s [c_j q^{-a_j+1}, (c_j+1)q^{-a_j+1})$ and where each $I \in \mathcal{P}_{\boldsymbol{c}}$ is of the form $I = \prod_{j=1}^s [x_j, y_j)$ with $q^{a_j}|y_j - x_j| \in \{1, \ldots, q-1\}$. We have $q^{-\sum_{j=1}^{s} a_j} \leq \mathrm{Vol}(I) \leq q^{-\sum_{j=1}^{s}(a_j-1)}$ and therefore

$$\frac{q^{-\sum_{j=1}^{s}(a_j-1)}}{\mathrm{Vol}(I)^{1-\lambda}} \leq \mathrm{Vol}(I)^{\lambda} q^s \leq q^{s-\lambda \sum_{j=1}^{s}(a_j-1)}$$

and hence

$$q^{-\sum_{j=1}^{s}(a_j-1)} \sup_J |\Delta(f, J)| \leq q^{s-\lambda \sum_{j=1}^{s}(a_j-1)} \sup_{\mathcal{P}_{\boldsymbol{c}}} \sum_{I \in \mathcal{P}_{\boldsymbol{c}}} \mathrm{Vol}(I)^{1-\lambda} |\Delta(f, I)|.$$

Note that

$$\sum_{\substack{0 \leq c_j < q^{a_j} - 1 \\ 1 \leq j \leq s}} \sup_{\mathcal{P}_{\boldsymbol{c}}} \sum_{I \in \mathcal{P}_{\boldsymbol{c}}} \mathrm{Vol}(I)^{1-\lambda} |\Delta(f, I)| \leq \sup_{\mathcal{P}} \sum_{J \in \mathcal{P}} \mathrm{Vol}(J) \frac{|\Delta(f, J)|}{\mathrm{Vol}(J)^{\lambda}}$$

where the supremum on the left hand side is taken over all partitions $\mathcal{P}$ of the cube $\prod_{j=1}^s [c_j q^{-a_j+1}, (c_j+1)q^{-a_j+1})$ into subintervals and the supremum on the right hand side is taken over all partitions of $[0,1)^s$ into subintervals. Thus the result follows. $\square$

For the next lemma we will need the following two functions. For $\kappa \in \{1, \ldots, q-1\}$ let now

$$v_\kappa = \sum_{r=0}^{q-1} r \mathrm{wal}_\kappa(r/q).$$

If q is chosen to be a prime number and the bijections φ and η are chosen to be the identity, then $v_\kappa = q(e^{2\pi i \kappa/q} - 1)^{-1}$, see [4, Appendix A].

Further for $l \in \{1, \ldots, q-1\}$ we define the function $\zeta_a(x) = \sum_{r=0}^{x_a-1} \overline{\mathrm{wal}_l(r/q)}$, where $a \geq 1$ and $x = x_1 q^{-1} + x_2 q^{-2} + \cdots$ and where for $x_a = 0$ we set $\zeta_a(x) = 0$. The function ζ_a depends on x only through x_a, thus it is a step-function which is constant on the intervals $[cq^{-a}, (c+1)q^{-a})$ for $c = 0, \ldots, q^a - 1$. By [25, Proposition 5] it follows that ζ_a can be represented by a finite Walsh series. Indeed, there are numbers $c_0, \ldots, c_{q-1}$ (which depend on l but not on a) such that

$$\zeta_a(x) = \sum_{z=0}^{q-1} c_z \overline{\mathrm{wal}_{zq^{a-1}}(x)}.$$

If q is chosen to be a prime number and the bijections φ and η are chosen to be the identity, then $\zeta_a(x) = (1 - \overline{\mathrm{wal}_{lq^{a-1}}(x)})(1 - \overline{\mathrm{wal}_l(1/q)})^{-1}$, i.e., $c_0 = (1 - \overline{\mathrm{wal}_l(1/q)})^{-1}$, $c_l = (\overline{\mathrm{wal}_l(1/q)} - 1)^{-1}$ and $c_z = 0$ for $z \neq 0, l$.

The following lemma will be used in the induction step for differentiable functions. For example, for a differentiable function $F : \mathbb{R} \to \mathbb{R}$ given by $F(x) = \int_0^x f(y) \, dy$ we can calculate the Walsh coefficients using integration by parts in the following way: for $k > 0$ we have

$$\hat{F}(k) = \int_0^1 F(x) \overline{\mathrm{wal}_k(x)} \, dx = \left[\int_0^x \overline{\mathrm{wal}_k(y)} \, dy F(x) \right]_0^1 - \int_0^1 f(x) \int_0^x \overline{\mathrm{wal}_k(y)} \, dy \, dx$$

$$= -\int_0^1 f(x) \int_0^x \overline{\mathrm{wal}_k(y)} \, dy \, dx, \tag{3.7}$$

where we used $\int_0^0 \overline{\mathrm{wal}_k(x)} \, dx = \int_0^1 \overline{\mathrm{wal}_k(x)} \, dx = 0$. For $k = 0$ on other hand we obtain

$$\hat{F}(0) = \int_0^1 F(x) \overline{\mathrm{wal}_0(x)} \, dx = \left[\int_0^x \overline{\mathrm{wal}_0(y)} \, dy F(x) \right]_0^1 - \int_0^1 f(x) \int_0^x \overline{\mathrm{wal}_0(y)} \, dy \, dx$$

$$= \int_0^1 f(x) \, dx - \int_0^1 f(x) \int_0^x \overline{\mathrm{wal}_0(y)} \, dy \, dx, \tag{3.8}$$

Thus if we know the Walsh series for f, then we can easily calculate the Walsh series for F, provided that we know the Walsh series for $\int_0^x \overline{\mathrm{wal}_k(y)} \, dy$. This will be calculated in the following lemma. It appeared in a simpler form in [10].

LEMMA 3.3. *For $k \in \mathbb{N}_0$ and $x \in [0,1)$ define $J_k(x) = \int_0^x \overline{\mathrm{wal}_k(y)} \, dy$. For $k \geq 1$ let $k = lq^{a-1} + k'$ where $l \in \{1, \ldots, q-1\}$, $a \geq 1$ and $0 \leq k' < q^{a-1}$. Then J_k can be represented by a Walsh series which is given by*

$$J_k(x) = q^{-a} \left(\sum_{z=0}^{q-1} c_z \overline{\mathrm{wal}_{zq^{a-1}+k'}(x)} + 2^{-1} \overline{\mathrm{wal}_k(x)} + \sum_{c=1}^{\infty} \sum_{\kappa=1}^{q-1} q^{-c-1} v_\kappa \overline{\mathrm{wal}_{\kappa q^{a+c-1}+k}(x)} \right).$$

Further we have

$$J_0(x) = 1/2 + \sum_{c=1}^{\infty} \sum_{\kappa=1}^{q-1} q^{-c-1} v_\kappa \overline{\mathrm{wal}_{\kappa q^{c-1}}(x)}.$$

Proof. Let $k = lq^{a-1} + k'$ with $a \geq 1$, $0 \leq k' < q^{a-1}$ and $l \in \{1, \ldots, q-1\}$. The function $\overline{\mathrm{wal}_{lq^{a-1}}(y)}$ is constant on each interval $[rq^{-a}, (r+1)q^{-a})$ and

$\overline{\mathrm{wal}_{k'}(y)}$ is constant on each interval $[cq^{-a+1}, (c+1)q^{-a+1})$. We have $\overline{\mathrm{wal}_k(y)} = \overline{\mathrm{wal}_{lq^{a-1}}(y)}\ \overline{\mathrm{wal}_{k'}(y)}$. For any $0 \le c < q^{a-1}$ we have

$$\int_{[cq^{-a+1},(c+1)q^{-a+1})} \overline{\mathrm{wal}_k(y)}\,\mathrm{d}y = \overline{\mathrm{wal}_{k'}(cq^{-a+1})}\int_{[cq^{-a+1},(c+1)q^{-a+1})} \overline{\mathrm{wal}_{lq^{a-1}}(y)}\,\mathrm{d}y$$

$$= \overline{\mathrm{wal}_{k'}(cq^{-a+1})}q^{-a}\sum_{r=0}^{q-1}\mathrm{wal}_l(r/q)$$

$$= 0.$$

Thus we have

$$J_k(x) = \overline{\mathrm{wal}_{k'}(x)}J_{lq^{a-1}}(x).$$

Let $x = x_1 q^{-1} + x_2 q^{-2} + \cdots$ and $y = x_{a+1}q^{-1} + x_{a+2}q^{-2} + \cdots$, then we have

$$J_{lq^{a-1}}(x) = q^{-a}\sum_{r=0}^{x_a-1}\overline{\mathrm{wal}_l(r/q)} + q^{-a}\overline{\mathrm{wal}_l(x_a/q)}y.$$

We now investigate the Walsh series representation of the function $J_{lq^{a-1}}(x)$. First note that $\overline{\mathrm{wal}_l(x_a/q)} = \overline{\mathrm{wal}_{lq^{a-1}}(x)}$. Further, by a slight adaption of [4, eq. (30)] we obtain

$$y = 1/2 + \sum_{c=1}^{\infty}\sum_{\kappa=1}^{q-1}q^{-c-1}v_\kappa\ \overline{\mathrm{wal}_{\kappa q^{c-1}}(y)}. \tag{3.9}$$

As $\overline{\mathrm{wal}_{\kappa q^{c-1}}(y)} = \overline{\mathrm{wal}_{\kappa q^{a+c-1}}(x)}$ we obtain

$$y = 1/2 + \sum_{c=1}^{\infty}\sum_{\kappa=1}^{q-1}q^{-c-1}v_\kappa\ \overline{\mathrm{wal}_{\kappa q^{a+c-1}}(x)}.$$

As noted above, the Walsh series of $\zeta_a(x) = \sum_{r=0}^{x_a-1}\overline{\mathrm{wal}_l(r/q)}$, where for $x_a = 0$ we set $\zeta_a(x) = 0$ can be written as

$$\zeta_a(x) = \sum_{z=0}^{q-1}c_z\overline{\mathrm{wal}_{zq^{a-1}}(x)}.$$

Altogether we obtain

$$q^a J_{lq^{a-1}}(x) = \sum_{z=0}^{q-1}c_z\overline{\mathrm{wal}_{zq^{a-1}}(x)} + 2^{-1}\overline{\mathrm{wal}_{lq^{a-1}}(x)}$$

$$+ \sum_{c=1}^{\infty}\sum_{\kappa=1}^{q-1}q^{-c-1}v_\kappa\overline{\mathrm{wal}_{\kappa q^{a+c-1}+lq^{a-1}}(x)}$$

and therefore

$$q^a J_k(x) = \sum_{z=0}^{q-1}c_z\overline{\mathrm{wal}_{zq^{a-1}+k'}(x)} + 2^{-1}\overline{\mathrm{wal}_k(x)} + \sum_{c=1}^{\infty}\sum_{\kappa=1}^{q-1}q^{-c-1}v_\kappa\overline{\mathrm{wal}_{\kappa q^{a+c-1}+k}(x)}.$$

The result for $k = 0$ follows easily from (3.9). □

Note that Lemma 3.3 can easily be generalized to arbitrary dimensions s, since for $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$ we have for any $\boldsymbol{x} = (x_1, \ldots, x_s) \in [0,1)^s$ that

$$J_{\boldsymbol{k}}(\boldsymbol{x}) = \int_{[0,\boldsymbol{x})} \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{y})}\, \mathrm{d}\boldsymbol{y} = \prod_{j=1}^{s} J_{k_j}(x_j),$$

where $[0, \boldsymbol{x}) = \prod_{j=1}^{s} [0, x_j)$.

The next lemma shows how the Walsh coefficients of a function $F = \int f$ can be obtained from the Walsh coefficients of f.

LEMMA 3.4. *Let $f \in \mathcal{L}_2([0,1)^s)$ and let $F(\boldsymbol{x}) = \int_{[0,\boldsymbol{x})} f(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}$, where $[0, \boldsymbol{x}) = \prod_{j=1}^{s}[0, x_j)$ with $\boldsymbol{x} = (x_1, \ldots, x_s)$. Further let $\hat{F}(\boldsymbol{k})$ denote the $\boldsymbol{k}$-th Walsh coefficient of F. Let $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$ and let $U = \{1 \le j \le s : k_j \ne 0\}$. For $j \in U$ let $k_j = l_j q^{a_j - 1} + k'_j$, $0 < l_j < q$ and $0 \le k'_j < q^{a_j - 1}$ and further let $\boldsymbol{k}' = (k'_1, \ldots, k'_s)$ where $k'_j = 0$ for $j \notin U$. Then we have*

$$\hat{F}(\boldsymbol{k}) = q^{-\sum_{j \in U} a_j} \sum_{U \subseteq v \subseteq \mathcal{S}} (-1)^{|v|} \sum_{\boldsymbol{h}_v \in \mathbb{N}_0^{|v|}} \hat{f}(\boldsymbol{k}' + (\boldsymbol{h}_v, \boldsymbol{0}))\, \chi_{U,v,\boldsymbol{k}}(\boldsymbol{h}_v),$$

where $(\boldsymbol{h}_v, \boldsymbol{0})$ denotes the s-dimensional vector whose j-th component is h_j for $j \in v$ and 0 otherwise and where for $\boldsymbol{h}_v = (h_j)_{j \in v} \in \mathbb{N}_0^{|v|}$ we set

$$\chi_{U,v,\boldsymbol{k}}(\boldsymbol{h}_v) = \prod_{j \in U} \rho_{k_j}(h_j) \prod_{j \in v \setminus U} \phi(h_j).$$

Here

$$\rho_{k_j}(h_j) = \begin{cases} c_z + 2^{-1}1_{z=l_j} & \text{for } h_j = z q^{a_j - 1}, \\ v_z q^{-i-1} & \text{for } h = z q^{a_j - 1 + i} + l_j q^{a_j - 1}, i > 0, 0 < z < q, \\ 0 & \text{otherwise}, \end{cases}$$

where $1_{z = l_j} = 1$ for $z = l_j$ and 0 otherwise, and

$$\phi(h_j) = \begin{cases} 2^{-1} & \text{for } h_j = 0, \\ v_z q^{-i-1} & \text{for } h = z q^{a_j - 1 + i}, i > 0, 0 < z < q, \\ 0 & \text{otherwise}. \end{cases}$$

Proof. Using integration by parts in each coordinate, Fubini's theorem and $J_k(0) = J_k(1) = 0$ for any $k \in \mathbb{N}$ (see Equations (3.7) and (3.8) for one-dimensional examples) it follows that

$$\hat{F}(\boldsymbol{k}) = \int_{[0,1)^s} F(\boldsymbol{x}) \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})}\, \mathrm{d}\boldsymbol{x} = \sum_{U \subseteq v \subseteq \mathcal{S}} (-1)^{|v|} \int_{[0,1)^s} J_{\boldsymbol{k}_v}(\boldsymbol{x}_v) f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x},$$

where for $\boldsymbol{k} = (k_1, \ldots, k_s)$ and $\boldsymbol{x} = (x_1, \ldots, x_s)$ we have $J_{\boldsymbol{k}_v}(\boldsymbol{x}_v) = \prod_{j \in v} J_{k_j}(x_j)$.

Using the Walsh series expansion of $J_{\boldsymbol{k}}$ given by Lemma 3.3 we can now express the Walsh coefficient $\hat{F}(\boldsymbol{k})$ as a sum of the Walsh coefficients $\hat{f}(\boldsymbol{h})$, from which the result follows. □

The following definition now captures the essence of the decay of the Walsh coefficients of smooth functions and will be used in the statement of the subsequent lemmas, theorems and corollaries.

DEFINITION 3.5. *Let* $k = k(v; a_1, \ldots, a_v) = \kappa_1 q^{a_1 - 1} + \cdots + \kappa_v q^{a_v - 1}$ *with* $v \geq 1$, $\kappa_1, \ldots, \kappa_v \in \{1, \ldots, q-1\}$ *and* $1 \leq a_v < \cdots < a_1$ *be a natural number. For* $k = 0$ *we set* $v = 0$, *i.e.,* $k(0) = 0$. *A function* $\mathcal{B} : \mathbb{N}_0 \to \mathbb{R}$ *is called* q*-adically non-increasing if* $\mathcal{B}(k) = \mathcal{B}(k(v; a_1, \ldots, a_v))$ *is non-increasing in* v *and each* a_i *for* $i = 1, \ldots, v$, *that is, for any* $v \geq 0$ *we have*

$$\mathcal{B}(k(v; a_1, \ldots, a_v)) \geq \mathcal{B}(k(v+1; a_1', \ldots, a_{v+1}'))$$

with $1 \leq a_{v+1}' < \cdots < a_1'$ *and* $a_1, \ldots, a_v \in \{a_1', \ldots, a_{v+1}'\}$ *and for an arbitrary* $1 \leq i \leq v$ *we have*

$$\mathcal{B}(k(v; a_1, \ldots, a_v)) \geq \mathcal{B}(k(v; a_1, \ldots, a_{i-1}, a_i + 1, a_{i+1}, \ldots, a_v))$$

provided that $a_i + 1 < a_{i-1}$ *in case* $1 < i \leq v$.

In the following lemma we give a bound on the Walsh coefficients of F if f satisfies some smoothness condition.

LEMMA 3.6. *Let* $\mathcal{B} : \mathbb{N}_0^s \to [0, \infty)$ *be a* q*-adically non-increasing function in each variable. Let* $f \in \mathcal{L}_2([0,1)^s)$ *and let the Walsh coefficients of* f *satisfy*

$$|\hat{f}(\boldsymbol{k})| \leq \mathcal{B}(\boldsymbol{k}) \quad \text{for all } \boldsymbol{k} \in \mathbb{N}_0^s.$$

Let $F(\boldsymbol{x}) = \int_{[0,\boldsymbol{x})} f(\boldsymbol{y}) \, d\boldsymbol{y}$, *where* $[0, \boldsymbol{x}) = \prod_{j=1}^s [0, x_j)$ *with* $\boldsymbol{x} = (x_1, \ldots, x_s)$. *Further let* $\hat{F}(\boldsymbol{k})$ *denote the* $\boldsymbol{k}$*-th Walsh coefficient of* F. *Let* $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}$ *and let* $U = \{1 \leq j \leq s : k_j \neq 0\}$. *For* $j \in U$ *let* $k_j = l_j q^{a_j - 1} + k_j'$ *and* $0 \leq k_j' < q^{a_j - 1}$ *and further let* $\boldsymbol{k}' = (k_1', \ldots, k_s')$ *where* $k_j' = 0$ *for* $j \notin U$. *Then there is a constant* $C_{s,U} > 0$ *independent of* $\boldsymbol{k}$ *such that*

$$|\hat{F}(\boldsymbol{k})| \leq C_{s,U} \, q^{-\sum_{j \in U} a_j} \mathcal{B}(\boldsymbol{k}').$$

Proof. Using Lemma 3.4 we obtain that

$$|\hat{F}(\boldsymbol{k})| \leq q^{-\sum_{j \in U} a_j} \mathcal{B}(\boldsymbol{k}') \sum_{U \subseteq v \subseteq \mathcal{S}} \sum_{\boldsymbol{h}_v \in \mathbb{N}_0^{|v|}} |\chi_{U,v,\boldsymbol{k}}(\boldsymbol{h}_v)|,$$

as $|\hat{f}(\boldsymbol{k}' + (\boldsymbol{h}_v, \boldsymbol{0}))| \leq \mathcal{B}(\boldsymbol{k}' + (\boldsymbol{h}_v, \boldsymbol{0})) \leq \mathcal{B}(\boldsymbol{k}')$ for all values of $\boldsymbol{h}_v \in \mathbb{N}_0^{|v|}$ for which $\chi_{U,v,\boldsymbol{k}}(\boldsymbol{h}_v) \neq 0$, since $\mathcal{B}$ is q-adically non-increasing in each variable.

Thus it remains to bound $\sum_{U \subseteq v \subseteq \mathcal{S}} \sum_{\boldsymbol{h}_v \in \mathbb{N}_0^{|v|}} |\chi_{U,v,\boldsymbol{k}}(\boldsymbol{h}_v)|$ independently of $\boldsymbol{k}$. We only prove the case where q is chosen to be a prime number and the bijections φ and η are chosen to be the identity, as in this case we can obtain an explicit constant $C_{s,U} > 0$. The general case can be obtained by similar arguments using the result from Lemma 3.4.

Using the notation from Lemma 3.4 we have

$$\sum_{h \in \mathbb{N}_0} |\rho_{k_j}(h)| = |1 - \omega_q^{-l_j}|^{-1} + 2^{-1}|1 + \omega_q^{-l_j}||\omega_q^{-l_j} - 1|^{-1}$$

$$+ \sum_{i=1}^{\infty} q^{-i} \sum_{z=1}^{q-1} |e^{2\pi i z/q} - 1|^{-1}$$

$$\leq 3(2 - 2\cos(2\pi/q))^{-1/2}$$

and

$$\sum_{h=0}^{\infty} |\phi(h)| = 2^{-1} + \sum_{i=1}^{\infty} q^{-i} \sum_{z=1}^{q-1} |e^{2\pi i z/q} - 1|^{-1} \leq 2^{-1} + (2 - 2\cos(2\pi/q))^{-1/2}.$$

Therefore we have

$$\sum_{U \subseteq v \subseteq S} \sum_{\boldsymbol{h}_v \in \mathbb{N}_0^{|v|}} |\chi_{U,v,\boldsymbol{k}}(\boldsymbol{h}_v)|$$

$$\leq 3^{|U|}(2 - 2\cos(2\pi/q))^{-|U|/2}(3/2 + (2 - 2\cos(2\pi/q))^{-1/2})^{s-|U|}$$

and hence we can choose

$$C_{s,U} = 3^{|U|}(2 - 2\cos(2\pi/q))^{-|U|/2}(3/2 + (2 - 2\cos(2\pi/q))^{-1/2})^{s-|U|} \qquad (3.10)$$

in Lemma 3.6 for this case. $\square$

We use the above results now to establish an upper bound on the Walsh coefficients of a polynomial. The proof will give a glimpse on how the argument will work for more general function classes.

LEMMA 3.7. *Let* $k = \kappa_1 q^{a_1-1} + \cdots + \kappa_v q^{a_v-1}$ *with* $v \geq 1$, $\kappa_1, \ldots, \kappa_v \in \{1, \ldots, q-1\}$ *and* $1 \leq a_v < \cdots < a_1$. *For* $v = 0$ *let* $k = 0$. *Let* $f : [0,1) \to \mathbb{R}$ *be the polynomial* $f(x) = f_0 + f_1 x + \cdots + f_i x^i$ *with* $f_i \neq 0$ *and let* $\hat{f}(k)$ *denote the* k-*th Walsh coefficient of* f. *Then for* $v \geq 0$ *there are constants* $0 < C_{f,i,v} < \infty$ *such that*

$$|\hat{f}(k)| \leq C_{f,i,v} q^{-a_1 - \cdots - a_v},$$

where we can choose $C_{f,i,v} = 0$ *for* $v > i$.

Proof. Let $f(x) = f_0 + f_1 x + \cdots + f_i x^i$, where $i = \deg(f)$ (that is, $f_i \neq 0$). Then we have $f^{(i)}(x) = i! f_i \neq 0$. As $f^{(i)}$ is a constant function, its Walsh series representation is simply given by $f^{(i)}(x) = i! f_i$. Now we use Lemma 3.6. The dimension s in our case is 1 and we can choose the function $\mathcal{B}_1$ by $\mathcal{B}_1(0) = i!|f_i|$ and for $k > 0$ we set $\mathcal{B}_1(k) = 0$. Note the function $\mathcal{B}_1$ defined this way is a q-adically non-increasing function. Then it follows that there is a constant $C_1 > 0$ such that the Walsh coefficients of the function $\int_0^x f^{(i)}(t) \, dt = f^{(i-1)}(x) - f^{(i-1)}(0)$ are bounded by $C_1 q^{-a_1} i! |f_i|$ for all k where $v = 1$ and the Walsh coefficients are 0 for $v > 1$. The Walsh coefficient for $k = 0$ is given by $f^{(i-2)}(1) - f^{(i-2)}(0) - f^{(i-1)}(0)$.

Now consider the function $\int_0^x f^{(i-1)}(t) \, dt = f^{(i-2)}(x) - f^{(i-2)}(0)$. It follows from the above and Lemma 3.6 that the Walsh coefficients of $f^{(i-2)}(x) - f^{(i-2)}(0)$ can be bounded by a q-adically non-increasing function $\mathcal{B}_2$. Indeed there are constants $C_2, C_3 > 0$ such that we can choose $\mathcal{B}_2(0) = |f^{(i-2)}(1) - f^{(i-2)}(0) - f^{(i-1)}(0)|$, $\mathcal{B}_2(k) = C_2 q^{-a_1}$ for $v = 1$, $\mathcal{B}_2(k) = C_3 q^{-a_1-a_2}$ for $v = 2$ and $\mathcal{B}_2(k) = 0$ for $v > 2$. Again $\mathcal{B}_2$ is a q-adically non-increasing function and Lemma 3.6 can again be used.

By using the above argument iteratively we obtain that there is a constant $C > 0$ such that $|\hat{f}(k)| \leq C q^{-a_1 - \cdots - a_v}$, for $k = \kappa_1 q^{a_1-1} + \cdots + \kappa_v q^{a_v-1}$ with $\kappa_1, \ldots, \kappa_v \in \{1, \ldots, q-1\}$ and $1 \leq a_v < \cdots < a_1$. The result thus follows. $\square$

For the case where q is chosen to be a prime number and the bijections φ and η are chosen to be the identity and for $0 \leq v \leq i$ we can choose

$$C_{f,i,v} = \bar{C}^v \sum_{l=v}^{i} C'^{i-l} l! |f_l|, \qquad (3.11)$$

where $\bar{C} = (2 - 2\cos(2\pi/q))^{-1/2}$ and $C' = 3/2 + (2 - 2\cos(2\pi/q))^{-1/2}$ in Lemma 3.7.

Let $f : [0,1)^s \to \mathbb{R}$ be such that the partial mixed derivatives up to order $\delta \geq 1$ in each variable exist and are continuous. We need some further notation: let $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_s)$ and

$$f^{(\boldsymbol{\tau})}(\boldsymbol{x}) = \frac{\partial^{\tau_1 + \cdots + \tau_s}}{\partial x_1^{\tau_1} \cdots \partial x_s^{\tau_s}} f(\boldsymbol{x}).$$

For $\boldsymbol{\tau} \in \{0, \ldots, \delta\}^s$ let $u(\boldsymbol{\tau}) = \{1 \leq j \leq s : \tau_j = \delta\}$. Let $\boldsymbol{\gamma} = (\gamma_v)_{v \subset \mathbb{N}}$ be an indexed set of non-negative real numbers. Let $v(\boldsymbol{\tau}) = \{1 \leq j \leq s : \tau_j > 0\}$. Then for $0 < \lambda \leq 1$ and $\mathfrak{p}, \mathfrak{q}, \mathfrak{r} \geq 1$ ($\mathfrak{p}, \mathfrak{q}, \mathfrak{r}$ do not appear in the subscript of N as they do not have influence on our subsequent bounds, we only assume that they are bigger or equal to 1) we define

$$N_{\delta, \lambda, \boldsymbol{\gamma}}(f) = \left(\sum_{\boldsymbol{\tau} \in \{0, \ldots, \delta\}^s} \gamma_{v(\boldsymbol{\tau})}^{-1} \left[V_{\lambda, \mathfrak{p}, \mathfrak{q}, \mathbf{1}}^{(|u(\boldsymbol{\tau})|)} (f^{(\boldsymbol{\tau})}(\cdot, \mathbf{0}_{\mathcal{S} \setminus u(\boldsymbol{\tau})})) \right]^{\mathfrak{r}} \right)^{1/\mathfrak{r}}, \qquad (3.12)$$

where, for clarity, we introduce the additional superscript $(|u(\boldsymbol{\tau})|)$ in the Hardy and Krause variation $V_{\lambda, \mathfrak{p}, \mathfrak{q}, \mathbf{1}}^{(|u(\boldsymbol{\tau})|)}$ which indicates the dimension of the function and where for $u(\boldsymbol{\tau}) = \emptyset$ we set $V_{\lambda, \mathfrak{p}, \mathfrak{q}, \mathbf{1}}^{(|u(\boldsymbol{\tau})|)} (f^{(\boldsymbol{\tau})}(\cdot, \mathbf{0}_{\mathcal{S} \setminus u(\boldsymbol{\tau})})) = |f^{(\boldsymbol{\tau})}(\mathbf{0})|$.

The weights $\boldsymbol{\gamma}$ are introduced to modify the importance of various coordinate projections and were first introduced in [30], see also [6, 7]. If for some $v' \subseteq \mathcal{S}$ the weight $\gamma_{v'} = 0$, then we assume that the function f satisfies $V_{\lambda, \mathfrak{p}, \mathfrak{q}, \mathbf{1}}^{(|v'|)} (f^{(\boldsymbol{\tau})}(\cdot, \mathbf{0}_{\mathcal{S} \setminus v'})) = 0$ for all $\boldsymbol{\tau} \in \{0, \ldots, \delta\}^s$ with $v(\boldsymbol{\tau}) = v'$ and in (3.12) we formally set $0/0 = 0$.

The parameters in the definition of $N_{\delta, \lambda, \boldsymbol{\gamma}}$ have the following meaning:

- δ denotes the order of partial derivatives of f required in order for $N_{\delta, \lambda, \boldsymbol{\gamma}}(f)$ to make sense;
- λ is a Hölder type parameter or fractional order type parameter of the generalized Hardy and Krause variation; roughly, f needs to have partial derivatives up to order $\delta + \lambda$, where for $0 < \lambda < 1$ this means some type of fractional smoothness or in dimension one a Hölder condition of order λ;
- the Vitali variation is in $\mathfrak{p}$ norm;
- $\mathfrak{q}$ is the norm in the summation of the generalized Hardy and Krause variation;
- $\mathfrak{r}$ is the norm in the summation over the $\boldsymbol{\tau}$;
- $\boldsymbol{\gamma}$ are the weights which regulate the importance of different coordinate projections;

Note that for $\lambda = 1$ and $\mathfrak{p} = \mathfrak{q} = \mathfrak{r} = 2$ the functional $N_{\delta, \lambda, \boldsymbol{\gamma}}$ is just the norm in a weighted reproducing kernel Sobolev space with continuous partial mixed derivatives up to order $\delta + 1$ in each variable. In one dimension the unweighted norm in this reproducing kernel Sobolev space is given by

$$\begin{aligned}
\langle f, g \rangle_{\mathrm{sob,an}, \delta+1} & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.13) \\
&= f(0)g(0) + \cdots + f^{(\delta-1)}(0)g^{(\delta-1)}(0) \\
&\quad + \int_0^1 f^{(\delta)}(x) \, \mathrm{d}x \int_0^1 g^{(\delta)}(x) \, \mathrm{d}x + \int_0^1 f^{(\delta+1)}(x)g^{(\delta+1)}(x) \, \mathrm{d}x
\end{aligned}$$

and for higher dimensions one just takes the weighted tensor product of the one dimensional reproducing kernel Sobolev spaces (see [7] for examples of weighted tensor

product reproducing kernel Sobolev spaces). Let the s dimensional weighted inner product be denoted by $\langle \cdot, \cdot \rangle_{\mathrm{sob,an},s,\delta+1,\boldsymbol{\gamma}}$ and the corresponding norm $\| \cdot \|_{\mathrm{sob,an},s,\delta+1,\boldsymbol{\gamma}}$ (indeed if the partial mixed derivatives up to order $\delta+1$ of f are continuous on $[0,1]^s$ then we have $\|f\|_{\mathrm{sob,an},s,\delta+1,\boldsymbol{\gamma}} = N_{\delta,1,\boldsymbol{\gamma}}(f)$).

In the following we define a function μ which will be used throughout the paper: let $\delta \geq 1$ be an integer and $0 < \lambda \leq 1$ be a real number. Then for $\boldsymbol{k} = (k_1, \ldots, k_s)$ we set

$$\mu_{q,\delta+\lambda}(\boldsymbol{k}) = \sum_{j=1}^{s} \mu_{q,\delta+\lambda}(k_j) \tag{3.14}$$

with

$$\mu_{q,\delta+\lambda}(k) = \begin{cases} 0 & \text{for } k = 0, \\ a_1 + \cdots + a_v & \text{for } v \leq \delta, \\ a_1 + \cdots + a_\delta + \lambda a_{\delta+1} & \text{for } v > \delta, \end{cases} \tag{3.15}$$

where for $k \in \mathbb{N}$ we write $k = \kappa_1 q^{a_1-1} + \cdots + \kappa_v q^{a_v-1}$ with $v \geq 1$, $\kappa_1, \ldots, \kappa_v \in \{1, \ldots, q-1\}$ and $1 \leq a_v < \cdots < a_1$.

THEOREM 3.8. *Let $\delta \geq 1$ be an integer, $0 < \lambda \leq 1$, $\mathfrak{p}, \mathfrak{q}, \mathfrak{r} \geq 1$ be real numbers and an indexed set $\boldsymbol{\gamma} = (\gamma_v)_{v \subset \mathbb{N}}$ of non-negative real numbers be given. Let $f : [0,1)^s \to \mathbb{R}$ be such that the partial mixed derivatives up to order δ in each variable exist and such that $N_{\delta,\lambda,\boldsymbol{\gamma}}(f) < \infty$. Then for any $\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}$ it follows that there is a constant $C_{f,q,s,\boldsymbol{\gamma}} > 0$ independent of $\boldsymbol{k}$ such that*

$$|\hat{f}(\boldsymbol{k})| \leq C_{f,q,s,\boldsymbol{\gamma}} q^{-\mu_{q,\delta+\lambda}(\boldsymbol{k})}.$$

Proof. In order to prove the result we use the Taylor series expansion of the function f. We have

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{\tau} \in \{0,\ldots,\delta-1\}^s} \frac{\boldsymbol{x}^{\boldsymbol{\tau}}}{\boldsymbol{\tau}!} f^{(\boldsymbol{\tau})}(\boldsymbol{0}) + \sum_{\emptyset \neq u \subseteq \mathcal{S}} \sum_{\boldsymbol{\tau}_{\mathcal{S}\setminus u} \in \{0,\ldots,\delta-1\}^{s-|u|}} ((\delta-1)!)^{-|u|}$$

$$\frac{\prod_{j \in \mathcal{S}\setminus u} x_j^{\tau_j}}{\prod_{j \in \mathcal{S}\setminus u} \tau_j!} \int_{[\boldsymbol{0}_u, \boldsymbol{x}_u)} f^{(\boldsymbol{\delta}_u, \boldsymbol{\tau}_{\mathcal{S}\setminus u})}(\boldsymbol{y}_u, \boldsymbol{0}_{\mathcal{S}\setminus u}) \prod_{j \in u} (x_j - y_j)^{\delta-1} \, \mathrm{d}\boldsymbol{y}_u. \tag{3.16}$$

First note that the first sum in (3.16) is a polynomial in $\boldsymbol{x}$ and therefore the Walsh coefficients of this polynomial satisfy the desired bound by Lemma 3.7.

Now we consider the second sum. Let $\emptyset \neq u \subseteq \mathcal{S}$ with $u = \{j_1, \ldots, j_{|u|}\}$ be given. Then for $j \notin u$ the Walsh coefficients satisfy the desired bound by Lemma 3.7. Hence it remains to consider the Walsh coefficients of

$$G_u(\boldsymbol{x}_u) = \int_{[\boldsymbol{0}_u, \boldsymbol{x}_u)} f^{(\boldsymbol{\delta}_u, \boldsymbol{\tau}_{\mathcal{S}\setminus u})}(\boldsymbol{y}_u, \boldsymbol{0}_{\mathcal{S}\setminus u}) \prod_{j \in u} (x_j - y_j)^{\delta-1} \, \mathrm{d}\boldsymbol{y}_u$$

$$= \int_{[\boldsymbol{0}_u, \boldsymbol{1}_u)} f^{(\boldsymbol{\delta}_u, \boldsymbol{\tau}_{\mathcal{S}\setminus u})}(\boldsymbol{y}_u, \boldsymbol{0}_{\mathcal{S}\setminus u}) \prod_{j \in u} (x_j - y_j)_+^{\delta-1} \, \mathrm{d}\boldsymbol{y}_u,$$

where $(x - y)_+ = \max(0, x - y)$.

By differentiating the function G_u in each variable $0 \leq k < \delta$ times we obtain (see [32, pp. 153,154])

$$\frac{\partial^{k|u|}}{\partial \boldsymbol{x}_u^k} G_u(\boldsymbol{x}_u)$$

$$= \left(\frac{(\delta-1)!}{(\delta-1-k)!} \right)^{|u|} \int_{[\boldsymbol{0}_u, \boldsymbol{x}_u)} f^{(\boldsymbol{\delta}_u, \boldsymbol{\tau}_{S \setminus u})}(\boldsymbol{y}_u, \boldsymbol{0}_{S \setminus u}) \prod_{j \in u} (x_j - y_j)^{\delta-1-k} \, \mathrm{d}\boldsymbol{y}_u.$$

Hence $\frac{\partial^{k|u|}}{\partial \boldsymbol{x}_u^k} G_u(\boldsymbol{x}_u) = 0$ if there is at least one $j \in u$ such that $x_j = 0$. Further we have

$$\frac{\partial^{\delta|u|}}{\partial \boldsymbol{x}_u^{\delta}} G_u(\boldsymbol{x}_u) = ((\delta-1)!)^{|u|} f^{(\boldsymbol{\delta}_u, \boldsymbol{\tau}_{S \setminus u})}(\boldsymbol{x}_u, \boldsymbol{0}_{S \setminus u}).$$

From $N_{\delta, \lambda, \boldsymbol{\gamma}}(f) < \infty$ it follows that the Walsh coefficients of $f^{(\boldsymbol{\delta}_u, \boldsymbol{\tau}_{S \setminus u})}(\boldsymbol{x}_u, \boldsymbol{0}_{S \setminus u})$ decay with order $\mu_{q, 0+\lambda}(k)$ in each variable. Further we have

$$G_u(\boldsymbol{x}_u) = \int_{[\boldsymbol{0}, \boldsymbol{x}_u)} \int_{[\boldsymbol{0}, \boldsymbol{y}_1)} \cdots \int_{[\boldsymbol{0}, \boldsymbol{y}_{\delta-1})} G_u^{(\boldsymbol{\delta}_u)}(\boldsymbol{y}_\delta) \, \mathrm{d}\boldsymbol{y}_\delta \cdots \mathrm{d}\boldsymbol{y}_1$$

as the function G_u and its derivatives are 0 if at least one $x_j = 0$ for $j \in u$, i.e., we have

$$\int_{[\boldsymbol{0}, \boldsymbol{x}_u)} G_u^{(\boldsymbol{\tau})}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = \sum_{v \subseteq u} (-1)^{u \setminus v} G_u^{(\boldsymbol{\tau}-\boldsymbol{1})}(\boldsymbol{x}_v, \boldsymbol{0}_{u \setminus v}) = G_u^{(\boldsymbol{\tau}-\boldsymbol{1})}(\boldsymbol{x}_u).$$

Hence it follows by repeated use of Lemma 3.6 that the desired bound holds for G_u and thus the result follows from (3.16). $\square$

For the case where q is chosen to be a prime number and the bijections φ and η are chosen to be the identity we can also obtain an explicit constant in Theorem 3.8. Indeed, using Lemma 3.2, Lemma 3.6 together with the explicit constant (3.10) and Lemma 3.7 together with the explicit constant (3.11) we obtain that the constant $C_{q, \boldsymbol{\gamma}}$ can be chosen as

$$C_{f, q, s, \boldsymbol{\gamma}} = \sum_{\substack{\boldsymbol{\tau} \in \{0, \ldots, \delta-1\}^s \\ \gamma_{v(\boldsymbol{\tau})} \neq 0}} |f^{(\boldsymbol{\tau})}(\boldsymbol{0})| \hat{C}^{\tau_1 + \cdots + \tau_s} + \sum_{\emptyset \neq u \subseteq S} q^{|u|} C_{s, u}^{\delta}$$

$$\sum_{\substack{\boldsymbol{\tau}_{S \setminus u} \in \{0, \ldots, \delta-1\}^{s-|u|} \\ \gamma_{u \cup v(\boldsymbol{\tau}_{S \setminus u})} \neq 0}} \hat{C}^{\sum_{j \in S \setminus u} \tau_j} V_{\lambda, 1}^{(|u|)}(f^{(\boldsymbol{\delta}_u, \boldsymbol{\tau}_{S \setminus u})}(\cdot, \boldsymbol{0}_{S \setminus u})),$$

where $\hat{C} = 1$ for $2 \leq q < 6$ and $\hat{C} = (2 - 2\cos(2\pi/q))^{-1/2}$ for $q > 6$ (note that for $q > 6$ we have $\hat{C} > 1$) and

$$C_{s, u} = 3^{|u|} (2 - 2\cos(2\pi/q))^{-|u|/2} (3/2 + (2 - 2\cos(2\pi/q))^{-1/2})^{s-|u|}.$$

As noted above, under certain conditions we can write $V_{\lambda, 1}^{(|u|)}$ also as an integral, see (3.4).

We give a further useful estimation of the constant by separating the dependence of the function from the constants. This way we obtain

$$C_{f, q, s, \boldsymbol{\gamma}} \leq C_{\delta, q, s, \boldsymbol{\gamma}} N_{\delta, \lambda, \boldsymbol{\gamma}}(f),$$

where

$$C_{\delta,q,s,\boldsymbol{\gamma}} = \sum_{\boldsymbol{\tau} \in \{0,\ldots,\delta-1\}^s} \gamma_{v(\boldsymbol{\tau})} \hat{C}^{\tau_1 + \cdots + \tau_s}$$
$$+ \sum_{\emptyset \neq u \subseteq \mathcal{S}} q^{|u|} C_{s,u}^\delta \sum_{\boldsymbol{\tau}_{\mathcal{S} \setminus u} \in \{0,\ldots,\delta-1\}^{s-|u|}} \gamma_{u \cup v(\boldsymbol{\tau}_{\mathcal{S} \setminus u})} \hat{C}^{\sum_{j \in \mathcal{S} \setminus u} \tau_j}. \quad (3.17)$$

Consider now the case where the weights are of product form (see [7]), i.e., there is a sequence of positive real numbers $(\gamma_j)_{j \in \mathbb{N}}$ such that $\gamma_v = \prod_{j \in v} \gamma_j$ for all $v \subset \mathbb{N}$ and for $v = \emptyset$ we set $\gamma_v = 1$. If now q is prime with $2 \leq q < 6$, then

$$C_{\delta,q,s,\boldsymbol{\gamma}} = \prod_{j=1}^s (1 + \gamma_j(\delta - 1)) + \prod_{j=1}^s \left[(1 + \gamma_j(\delta - 1))(3/2 + \hat{C}) + \gamma_j 3q\hat{C} \right]$$
$$- \prod_{j=1}^s \left[(1 + \gamma_j(\delta - 1))(3/2 + \hat{C}) \right].$$

For example for $q = 2$, $\delta = 1$ and product weights we obtain

$$C_{1,2,s,\boldsymbol{\gamma}} = 1 - 2^s + 2^s \prod_{j=1}^s (1 + 6\gamma_j).$$

The approach used here for prime q and φ the identity map can also be used for arbitrary prime powers q and arbitrary mappings φ with $\varphi(0) = 0$. Hence we obtain the following corollary.

COROLLARY 3.9. *Under the assumptions of Theorem 3.8 there exists a constant* $C_{\delta,q,s,\boldsymbol{\gamma}} > 0$ *independent of* $\boldsymbol{k}$ *and* f *such that*

$$|\hat{f}(\boldsymbol{k})| \leq C_{\delta,q,s,\boldsymbol{\gamma}} N_{\delta,\lambda,\boldsymbol{\gamma}}(f) q^{-\mu_{q,\delta+\lambda}(\boldsymbol{k})} \quad \text{for all } \boldsymbol{k} \in \mathbb{N}_0^s,$$

where $\mu_{q,\delta+\lambda}$ *is given by (3.14) and (3.15).*

REMARK 3.10. The results in this section also hold for the following generalization. In the definition of $N_{\delta,\lambda,\boldsymbol{\gamma}}$ we anchored the function and its derivatives at 0, i.e., we used $V_{\lambda,\mathbf{p},\mathbf{q},\mathbf{1}}^{(|u(\boldsymbol{\tau})|)}(f(\cdot, \mathbf{0}_{\mathcal{S} \setminus u(\boldsymbol{\tau})}))$. This can be generalized by choosing an arbitrary $\boldsymbol{a} \in [0,1]^s$ and using $V_{\lambda,\mathbf{p},\mathbf{q},\mathbf{1}}^{(|u(\boldsymbol{\tau})|)}(f(\cdot, \boldsymbol{a}_{\mathcal{S} \setminus u(\boldsymbol{\tau})}))$ in the definition of $N_{\delta,\lambda,\boldsymbol{\gamma}}$. It can be shown that in this case we also have Theorem 3.8 and Corollary 3.9.

3.3. Convergence of the Walsh series. For our purposes here we need strong assumptions on the convergence of the Walsh series $S(f)(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} \hat{f}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})$ to the function f, i.e., we require that the partial series $S_m(f)(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \\ k_j < m}} \hat{f}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})$ converges to $f(\boldsymbol{x})$ at every point $\boldsymbol{x} \in [0,1)^s$ as $m \to \infty$. (Note that the Walsh series $S(f)$ for the functions considered in this paper is always absolutely convergent, i.e., $\sum_{\boldsymbol{k} \in \mathbb{N}_0^s} |\hat{f}(\boldsymbol{k})| < \infty$, hence the Walsh series $S(f)(\boldsymbol{x})$ is uniformly bounded by $\sum_{\boldsymbol{k} \in \mathbb{N}_0^s} |\hat{f}(\boldsymbol{k})|$ and therefore $S(f)(\boldsymbol{x})$ itself converges at every point $\boldsymbol{x} \in [0,1)^s$.) This is necessary as we want to approximate the integral at function values $\boldsymbol{x}_n$ and for our analysis we deal with the Walsh series rather than the function itself, hence it is paramount that the function and its Walsh series coincide at every point $\boldsymbol{x} \in [0,1)^s$.

As the functions considered here are at least differentiable it follows that they are continuous and using the argument in [10, p. 373] it follows that the Walsh series

really converges at every point $\boldsymbol{x} \in [0,1)^s$ to the function value $f(\boldsymbol{x})$. Indeed, for a given $\boldsymbol{x} \in [0,1)^s$ we have

$$S_{q^m}(f)(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \{0,\ldots,q^m-1\}^s} \hat{f}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) = \mathrm{Vol}(J_{\boldsymbol{x}})^{-1} \int_{J_{\boldsymbol{x}}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x},$$

where $J_{\boldsymbol{x}} = \prod_{j=1}^{s} [q^{-m} \lfloor q^m x_j \rfloor, q^{-m} \lfloor q^m x_j \rfloor + q^{-m})$. The last equality follows from

$$\sum_{\boldsymbol{k} \in \{0,\ldots,q^m-1\}^s} \hat{f}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})$$

$$= \int_{[0,1)^s} f(\boldsymbol{y}) \sum_{\boldsymbol{k} \in \{0,\ldots,q^m-1\}^s} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{y})} \, \mathrm{d}\boldsymbol{y}$$

$$= \mathrm{Vol}(J_{\boldsymbol{x}})^{-1} \int_{J_{\boldsymbol{x}}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

As the function f is continuous it immediately follows that $S_{q^m}(f)(\boldsymbol{x})$ converges to $f(\boldsymbol{x})$ as m goes to infinity and the result follows.

3.4. A function space based on Walsh functions containing smooth functions. In this section we use the above results to define a function space based on Walsh functions which contains smooth functions for smoothness conditions considered in the previous section.

Let $\vartheta > 1$ be a real number and q a prime power. Then for $\boldsymbol{k} \in \mathbb{N}_0^s$ we set $r_{q,\vartheta}(\boldsymbol{k}) = q^{-\mu_{q,\vartheta}(\boldsymbol{k})}$, where $\mu_{q,\vartheta}$ is given by (3.14) and (3.15) (if ϑ is an integer, then choose $\lambda = 1$ and $\delta = \vartheta - 1$ and otherwise $\delta = \lfloor \vartheta \rfloor$ and $\lambda = \vartheta - \lfloor \vartheta \rfloor$).

Now we define a function space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}} \subseteq \mathcal{L}_2([0,1)^s)$ with norm $\|\cdot\|_{\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}}$ given by

$$\|f\|_{\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}} = \max_{\substack{u \subseteq \mathcal{S} \\ \gamma_u \neq 0}} \gamma_u^{-1} \sup_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} \frac{|\hat{f}(\boldsymbol{k}_u, \boldsymbol{0}_{\mathcal{S} \setminus u})|}{r_{q,\vartheta}(\boldsymbol{k}_u)},$$

where again for $\gamma_u = 0$ we assume that $\hat{f}(\boldsymbol{k}_u, \boldsymbol{0}_{\mathcal{S} \setminus u}) = 0$ for all $\boldsymbol{k}_u \in \mathbb{N}^{|u|}$.

The following result follows now directly from Corollary 3.9.

COROLLARY 3.11. *Let $\delta \geq 1$, $0 < \lambda \leq 1$, $\mathfrak{p}, \mathfrak{q}, \mathfrak{r} \geq 1$ and an indexed set $\boldsymbol{\gamma} = (\gamma_v)_{v \subset \mathbb{N}}$ of non-negative real numbers be given. Then there exists a constant $C_{\delta,q,s,\boldsymbol{\gamma}} > 0$ such that for every function $f : [0,1)^s \to \mathbb{R}$, whose partial mixed derivatives up to order δ exist, we have*

$$\|f\|_{\mathcal{E}_{s,q,\delta+\lambda,\boldsymbol{\gamma}}} \leq C_{\delta,q,s,\boldsymbol{\gamma}} N_{\delta,\lambda,\boldsymbol{\gamma}}(f),$$

where $\mu_{q,\delta+\lambda}$ is given by (3.14) and (3.15).

Again, using (3.17) an explicit constant in Corollary 3.11 can be obtained for q prime and φ the identity map. For all other cases (i.e., arbitrary prime powers q and mappings φ with $\varphi(0) = 0$) explicit constants can be obtained as well, but in this case the constant may also depend on the particular choice of q and φ. Further, as noted already above, for $\lambda = 1$ and $\mathfrak{p} = \mathfrak{q} = \mathfrak{r} = 2$ the functional $N_{\delta,\lambda,\boldsymbol{\gamma}}$ coincides with the norm in a certain Sobolev space (the one dimensional inner product for this Sobolev space is given by (3.13) and for higher dimensions one just considers tensor products of the one dimensional space) and hence it follows that $\mathcal{E}_{s,q,\delta+1,\boldsymbol{\gamma}}$ contains

certain Sobolev spaces. Hence Corollary 3.11 shows that if we want to prove results for smooth functions it is enough to consider $\mathcal{E}_{s,q,\vartheta,\gamma}$ (in the following we design quasi-Monte Carlo rules which work well for $\mathcal{E}_{s,q,\vartheta,\gamma}$ rather than directly for smooth functions, so the results for smooth functions come as a byproduct).

A function $f \in \mathcal{E}_{s,q,\vartheta,\gamma}$ can be written as a sum of their anova terms $f = \sum_{u \subseteq \mathcal{S}} f_u$ (see [8]). For a function $f \in \mathcal{E}_{s,q,\vartheta,\gamma}$ given by $f(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} \hat{f}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})$ the anova term f_u corresponding to a subset $u \subseteq \mathcal{S}$ is simply given by

$$f_u(\boldsymbol{x}_u) = \sum_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} \hat{f}(\boldsymbol{k}_u, \boldsymbol{0}_{\mathcal{S} \setminus u}) \mathrm{wal}_{\boldsymbol{k}_u}(\boldsymbol{x}_u).$$

If for some $u \subseteq \mathcal{S}$ we have $\gamma_u = 0$, then this implies that the anova term corresponding to u satisfies $f_u \equiv 0$. Hence the Walsh space $\mathcal{E}_{s,q,\vartheta,\gamma}$ consists only of functions whose anova term belonging to a subset u is zero for all subsets u with $\gamma_u = 0$ (see also [7]).

4. Digital $(t, \alpha, \beta, n \times m, s)$-nets and digital $(t, \alpha, \beta, \sigma, s)$-sequences. In this section we give the definition of digital $(t, \alpha, \beta, n \times m, s)$-nets and digital $(t, \alpha, \beta, \sigma, s)$-sequences. Similar point sets were introduced in [2].

4.1. The digital construction scheme. The construction of the point set used here is a slight generalization of the digital construction scheme introduced by Niederreiter, see [20], by breaking with the tradition of having square generating matrices.

DEFINITION 4.1. Let q be a prime-power and let $n, m, s \geq 1$ be integers. Let $C_1, \ldots, C_s$ be $n \times m$ matrices over the finite field $\mathbb{F}_q$ of order q. Now we construct q^m points in $[0, 1)^s$: for $0 \leq h \leq q^m - 1$ let $h = h_0 + h_1 q + \cdots + h_{m-1} q^{m-1}$ be the q-adic expansion of h. Consider an arbitrary but fixed bijection $\varphi : \{0, 1, \ldots, q-1\} \longrightarrow \mathbb{F}_q$. Identify h with the vector $\vec{h} = (\varphi(h_0), \ldots, \varphi(h_{m-1}))^\top \in \mathbb{F}_q^m$, where $\top$ means the transpose of the vector (note that we write $\vec{h}$ for vectors in the finite field $\mathbb{F}_q^m$ and $\boldsymbol{h}$ for vectors of integers or real numbers). For $1 \leq j \leq s$ multiply the matrix C_j by $\vec{h}$, i.e.,

$$C_j \vec{h} =: (y_{j,1}(h), \ldots, y_{j,n}(h))^\top \in \mathbb{F}_q^n,$$

and set

$$x_{h,j} := \frac{\varphi^{-1}(y_{j,1}(h))}{q} + \cdots + \frac{\varphi^{-1}(y_{j,n}(h))}{q^n}.$$

The point set $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{q^m-1}\}$ is called a digital net (over $\mathbb{F}_q$) (with generating matrices $C_1, \ldots, C_s$).

For $n, m = \infty$ we obtain a sequence $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots\}$, which is called a digital sequence (over $\mathbb{F}_q$) (with generating matrices $C_1, \ldots, C_s$).

Niederreiter's concept of a digital (t, m, s)-net and a digital (t, s)-sequence will appear as a special case in the subsequent section. Further, the digital nets considered below all satisfy $n \geq m$.

For a digital net with generating matrices $C_1, \ldots, C_s$ let $\mathcal{D} = \mathcal{D}(C_1, \ldots, C_s)$ be the dual net given by

$$\mathcal{D} = \{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} : C_1^\top \vec{k}_1 + \cdots + C_s^\top \vec{k}_s = \vec{0}\},$$

where for $\boldsymbol{k} = (k_1, \ldots, k_s)$ with $k_j = \kappa_{j,0} + \kappa_{j,1} q + \cdots$ and $\kappa_{j,i} \in \{0, \ldots, q-1\}$ let $\vec{k}_j = (\varphi(\kappa_{j,0}), \ldots, \varphi(\kappa_{j,n-1}))^\top$. Further, for $\emptyset \neq u \subseteq \mathcal{S}$ let $\mathcal{D}_u = \mathcal{D}((C_j)_{j \in u})$ and $\mathcal{D}_u^* = \mathcal{D}_u \cap \mathbb{N}^{|u|}$.

Note that throughout the paper Walsh functions and digital nets are defined using the same finite field $\mathbb{F}_q$ and the same bijection φ.

The following lemma is a slight generalization of [26, Lemma 2.5].

LEMMA 4.2. *Let* $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{q^m-1}\}$ *be a digital net over* $\mathbb{F}_q$ *with bijection* φ*, where* $\varphi(0) = 0$*, generated by the* $n \times m$ *matrices* $C_1, \ldots, C_s$ *over* $\mathbb{F}_q$*,* $n, m \geq 1$*. Then for any vector* $\boldsymbol{k} = (k_1, \ldots, k_s)$ *of nonnegative integers* $0 \leq k_1, \ldots, k_s < q^n$ *we have*

$$\sum_{h=0}^{q^m-1} {}_{\mathbb{F}_q,\varphi}\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_h) = \begin{cases} q^m & \text{if } \boldsymbol{k} \in \mathcal{D} \cup \{\boldsymbol{0}\}, \\ 0 & \text{else,} \end{cases}$$

where $\boldsymbol{0}$ *is the zero vector in* $\mathbb{N}_0^s$*.*

4.2. $(t, \alpha, \beta, n \times m, s)$-nets and $(t, \alpha, \beta, \sigma, s)$-sequences. Digital (t, α, β, m, s)-nets and digital (t, α, β, s)-sequences were first introduced in [2]. Those point sets were used for quasi-Monte Carlo rules which achieve the optimal rate of convergence of the worst-case error in Korobov spaces (which are reproducing kernel Hilbert spaces of smooth periodic functions). By a slight generalization of digital (t, α, β, m, s)-nets we will show that those digital nets also achieve the optimal convergence of the worst-case error in the space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ for all $1 < \vartheta \leq \alpha$.

The t value of a (t, m, s)-net is a quality parameter for the distribution properties of the net. A low t value yields well distributed point sets and it has been shown, see for example [4, 20], that a small t value also guarantees a small worst-case error for integration in Sobolev spaces for which the partial first derivatives are square integrable. In [2] it was shown how the definition of the t value needs to be modified in order to obtain faster convergence rates for periodic Sobolev spaces for which the partial derivatives up to order $\delta \leq \beta$ are square integrable. Here we extend those result in several ways. First we generalize the digital (t, α, β, m, s)-nets used in [2] to digital $(t, \alpha, \beta, n \times m, s)$-nets and show that we then can remove the periodicity assumption necessary in [2]. Further, if the derivatives up to order δ also have bounded variation with coefficient $0 < \lambda \leq 1$, then we have shown that such functions are in $\mathcal{E}_{s,q,\delta+\lambda,\boldsymbol{\gamma}}$.

In the following we repeat some definitions and results from [2] and give the definition of digital $(t, \alpha, \beta, n \times m, s)$-nets and digital $(t, \alpha, \beta, \sigma, s)$-sequences.

For a real number $\vartheta > 1$ the definition of the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ suggests to define the metric $\mu_{q,\vartheta}(\boldsymbol{k}, \boldsymbol{l}) = \mu_{q,\vartheta}(\boldsymbol{k} \ominus \boldsymbol{l})$ on $\mathbb{N}_0^s$, where $\mu_{q,\vartheta}(\boldsymbol{k} \ominus \boldsymbol{l})$ is given by (3.14) and (3.15), which is an extension of the metric introduced in [19], see also [28] (the metric for $\vartheta = 1$ can be used for Walsh spaces for example considered in [4]; for this case one basically obtains the metric in [19, 28]). As we will see later, in order to obtain a small worst-case error in the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ we need digital nets for which $\min\{\mu_{q,\vartheta}(\boldsymbol{k}) : \boldsymbol{k} \in \mathcal{D}\}$ is large. By translating this property into a linear independence property of the row vectors of the generating matrices $C_1, \ldots, C_s$ we arrive at the following definition.

DEFINITION 4.3. Let $n, m, \alpha \geq 1$ be natural numbers, let $0 < \beta \leq \alpha m/n$ be a real number and let $0 \leq t \leq \beta n$ be a natural number. Let $\mathbb{F}_q$ be the finite field of prime power order q and let $C_1, \ldots, C_s \in \mathbb{F}_q^{n \times m}$ with $C_j = (c_{j,1}, \ldots, c_{j,n})^\top$. If for all

$1 \leq i_{j,\nu_j} < \cdots < i_{j,1} \leq n$, where $0 \leq \nu_j \leq m$ for all $j = 1, \ldots, s$, with

$$\sum_{j=1}^{s} \sum_{l=1}^{\min(\nu_j, \alpha)} i_{j,l} \leq \beta n - t$$

the vectors

$$c_{1,i_{1,\nu_1}}, \ldots, c_{1,i_{1,1}}, \ldots, c_{s,i_{s,\nu_s}}, \ldots, c_{s,i_{s,1}}$$

are linearly independent over $\mathbb{F}_q$ then the digital net with generating matrices $C_1, \ldots, C_s$ is called a digital $(t, \alpha, \beta, n \times m, s)$-net over $\mathbb{F}_q$. Further we call a digital $(t, \alpha, \beta, n \times m, s)$-net over $\mathbb{F}_q$ with the largest possible value of β, i.e., $\beta = \alpha m/n$, a digital $(t, \alpha, n \times m, s)$-net over $\mathbb{F}_q$.

If t is the smallest non-negative integer such that the digital net generated by $C_1, \ldots, C_s$ is a digital $(t, \alpha, \beta, n \times m, s)$-net, then we call the digital net a strict digital $(t, \alpha, \beta, n \times m, s)$-net or a strict digital $(t, \alpha, n \times m, s)$-net if $\beta = \alpha m/n$.

REMARK 4.4. Using duality theory (see [22]) it follows that for a digital $(t, \alpha, \beta, n \times m, s)$-net we have $\min_{\boldsymbol{k} \in \mathcal{D}} \mu_{q,\alpha}(\boldsymbol{k}) > \beta n - t$ and for a strict digital $(t, \alpha, \beta, n \times m, s)$-net we have $\min_{\boldsymbol{k} \in \mathcal{D}} \mu_{q,\alpha}(\boldsymbol{k}) = \beta n - t + 1$. Hence digital $(t, \alpha, \beta, n \times m, s)$-nets with high quality have a large value of $\beta n - t$.

REMARK 4.5. In summary the parameters $t, \alpha, \beta, n, m, s$ have the following meaning:

- s denotes the dimension of the point set.
- n and m denote the size of the generating matrices for digital nets, i.e. the generating matrices are of size $n \times m$; in particular this means the point set has q^m points.
- t denotes the quality parameter of the point set; a low t value means high quality. In the upper bound, t is a quality parameter related to the constant in the upper bound.
- β is also a quality parameter. We will see later that the integration error is roughly q^{-n}. This is of course only true within boundaries, which is the reason for the parameter β, i.e. the integration error is roughly $q^{-\beta n}$. Hence β is a quality parameter related to the convergence rate.
- α is the smoothness parameter of the point set.

We can group the parameters also in the following way:

- m, n, s are fixed parameters, i.e. they specify the number and size of the generating matrices.
- α is a variable parameter, i.e. given (fixed) generating matrices can for example generate a $(t_1, 1, \beta_1, 10 \times 5, 5)$-net, a $(t_2, 2, \beta_2, 10 \times 5, 5)$-net, and so on (note the point set is always the same in each instance; the values $t_1, t_2, \ldots, \beta_1, \beta_2, \ldots$ may differ). This is necessary as in the upper bounds α will be the smoothness of the integrand, which may not be known explicitly.
- t and β are dependent parameters, they will depend on the generating matrices and on α. For given generating matrices, it is desirable to know the values of β and t for each value of $\alpha \in \mathbb{N}$.

Digital $(t, \alpha, \beta, n \times m, s)$-nets do not exist for arbitrary choices of the parameters $t, \alpha, \beta, n, m, s$, see [2]. The digital nets considered in [2] had the restriction that $n = m$ and special attention was paid to those digital nets with high quality, i.e., where $\alpha = \beta$. In this paper, a special role will be played by those digital nets for which $n = \alpha m$ and $\beta = 1$. The restriction on the linear independence of the digital nets comprises

now $n - t = \alpha m - t$ row vectors, which is the same as in [2], with the only difference that the size of the generating matrices is now bigger as now each generating matrix has $n = \alpha m$ rows. As those digital nets play a special role in this work we have the following definition.

DEFINITION 4.6.   A digital $(t, \alpha, 1, \alpha m \times m, s)$-net over $\mathbb{F}_q$ is called a digital $(t, \alpha, \alpha m \times m, s)$-net over $\mathbb{F}_q$. A strict digital $(t, \alpha, 1, \alpha m \times m, s)$-net over $\mathbb{F}_q$ is called a strict digital $(t, \alpha, \alpha m \times m, s)$-net over $\mathbb{F}_q$.

REMARK 4.7.   For practical purposes we would like to explicitly know digital $(t, \alpha, \alpha m \times m, s)$-nets for all $\alpha, m, s \geq 1$ with t as small as possible (as will be shown later, they achieve the optimal rate of convergence of the integration error of integrands for which all mixed partial derivatives of order α are, for example, square integrable, thus their usefulness).

Further, for given $\alpha, m, s \geq 1$ and a given digital $(t, \alpha, \alpha m \times m, s)$-net P, we would then also like to know the t' and β' value of this point set P when viewed as a digital $(t', \delta, \beta', \alpha m \times m, s)$-net for all values $\delta \in \mathbb{N}$, i.e., t' and β' are functions of δ (this is because we would also like to know how well such a digital net P performs if the integrand has partial mixed derivatives of order up to δ, because we might not know the smoothness of the integrand, but still would wish that P performs best possible).

We can also define sequences of points for which the first q^m points form a digital $(t, \alpha, \beta, n \times m, s)$-nets. In the classical case [20] one can just consider the left-upper $m \times m$ submatrices of the generating matrices of a digital sequence and determine the net properties of these for each $m \in \mathbb{N}$. Here, on the other hand, we are considering digital nets whose generating matrices are $n \times m$ matrices. So we would have to consider the left-upper $n_m \times m$ submatrices of the generating matrices of the digital sequence for each $m \in \mathbb{N}$ and where $(n_m)_{m \in \mathbb{N}}$ is a sequence of natural numbers. For our purposes here it is enough to consider only n_m of the form σm, for some given $\sigma \in \mathbb{N}$.

DEFINITION 4.8.   Let $\alpha, \sigma \geq 1$ and $t \geq 0$ be integers and let $0 < \beta \leq \alpha/\sigma$ be a real number. Let $\mathbb{F}_q$ be the finite field of prime power order q and let $C_1, \ldots, C_s \in \mathbb{F}_q^{\infty \times \infty}$ with $C_j = (c_{j,1}, c_{j,2}, \ldots)^\top$. Further let $C_{j, \sigma m \times m}$ denote the left upper $\sigma m \times m$ submatrix of C_j. If for all $m > t/(\beta \sigma)$ the matrices $C_{1, \sigma m \times m}, \ldots, C_{s, \sigma m \times m}$ generate a digital $(t, \alpha, \beta, \sigma m \times m, s)$-net then the digital sequence with generating matrices $C_1, \ldots, C_s$ is called a digital $(t, \alpha, \beta, \sigma, s)$-sequence over $\mathbb{F}_q$. Further we call a digital $(t, \alpha, 1, \alpha, s)$-sequence over $\mathbb{F}_q$ a digital (t, α, s)-sequence over $\mathbb{F}_q$.

If t is the smallest non-negative integer such that the digital sequence generated by $C_1, \ldots, C_s$ is a digital $(t, \alpha, \beta, \sigma, s)$-sequence, then we call the digital sequence a strict digital $(t, \alpha, \beta, \sigma, s)$-sequence or a strict digital (t, α, s)-sequence if $\alpha = \sigma$ and $\beta = 1$.

For short we will often write $(t, \alpha, \beta, n \times m, s)$-net instead of digital $(t, \alpha, \beta, n \times m, s)$-net over $\mathbb{F}_q$. The same applies to the other notions defined above.

REMARK 4.9.   Note that the definition of a digital $(t, 1, m \times m, s)$-net coincides with the definition of a digital (t, m, s)-net and the definition of a digital $(t, 1, s)$-sequence coincides with the definition of a digital (t, s)-sequence as defined by Niederreiter [20]. Further note that the t-value depends on α, β and σ, i.e., $t = t(\alpha, \beta, \sigma)$ or $t = t(\alpha)$ if $\alpha = \sigma$ and $\beta = 1$.

The definition of (t, α, s)-sequences here differs slightly from the definition in [2]. Indeed the definition of a (t, α, s)-sequence in [2] corresponds to a $(t, \alpha, \alpha, 1, s)$-sequence in the terminology of this paper, whereas here we call a $(t, \alpha, 1, \alpha, s)$-sequence a (t, α, s)-sequence. On the other hand note that the condition of linear independence

in Definition 4.3 is the same in both cases, i.e., the sum $i_{1,1} + \cdots + i_{1,\min(\nu_1,\alpha)} + \cdots + i_{s,1} + \cdots + i_{s,\min(\nu_s,\alpha)}$ needs to be bounded by $\alpha m - t$ for all m for $(t, \alpha, 1, \alpha, s)$-sequences and also for $(t, \alpha, \alpha, 1, s)$-sequences.

4.3. Some properties of $(t, \alpha, \beta, n \times m, s)$-nets and $(t, \alpha, \beta, \sigma, s)$-sequences. The properties of such digital nets and sequences shown in [2] also hold here. For example it was shown there that a digital (t, α, m, s)-net is also a digital $(\lceil t\alpha'/\alpha \rceil, \alpha', m, s)$-net for all $1 \le \alpha' \le \alpha$ and every digital (t, α, s)-sequence is also a digital $(\lceil t\alpha'/\alpha \rceil, \alpha', s)$-sequence for all $1 \le \alpha' \le \alpha$. In the same way we have the following theorem.

THEOREM 4.10. *Let P be a digital $(t, \alpha, \beta, n \times m, s)$-net over $\mathbb{F}_q$ and let S be a digital $(t, \alpha, \beta, \sigma, s)$-sequence over $\mathbb{F}_q$. Then we have:*

(i) *P is a digital $(t', \alpha, \beta', n \times m, s)$-net for all $1 \le \beta' \le \beta$ and all $t \le t' \le \beta' m$ and S is a digital $(t', \alpha, \beta', \sigma, s)$-sequence for all $1 \le \beta' \le \beta$ and all $t \le t'$.*

(ii) *P is a digital $(t', \alpha', \beta', n \times m, s)$-net for all $1 \le \alpha' \le n$ where $\beta' = \beta \min(\alpha, \alpha')/\alpha$ and $t' = \lceil t \min(\alpha, \alpha')/\alpha \rceil$ and S is a digital $(t', \alpha', \beta', \sigma, s)$-sequence for all $\alpha' \ge 1$ where $\beta' = \beta \min(\alpha, \alpha')/\alpha$ and where $t' = \lceil t \min(\alpha, \alpha')/\alpha \rceil$.*

(iii) *Any digital $(t, \alpha, n \times m, s)$-net is a digital $(\lceil t\alpha'/\alpha \rceil, \alpha', n \times m, s)$-net for all $1 \le \alpha' \le \alpha$ and every digital (t, α, σ, s)-sequence is a digital $(\lceil t\alpha'/\alpha \rceil, \alpha', \sigma, s)$-sequence for all $1 \le \alpha' \le \alpha$.*

(iv) *If $C_1, \ldots, C_s \in \mathbb{Z}_b^{n \times m}$ are the generating matrices of a digital $(t, \alpha, \beta, n \times m, s)$-net then the matrices $C_1^{(n')}, \ldots, C_s^{(n')}$, where $C_j^{(n')}$ consists of the first n' rows of C_j, generate a digital $(t, \alpha, \beta, n' \times m, s)$-net for all $1 \le n' \le n$.*

(v) *Any digital $(t, \alpha, \beta, \sigma, s)$-sequence is a digital $(t, \alpha, \beta, \sigma', s)$-sequence for all $1 \le \sigma' \le \sigma$.*

4.4. Constructions of $(t, \alpha, \beta, n \times m, s)$-nets and (t, α, σ, s)-sequences. In this section we show how explicit examples of $(t, \alpha, \beta, n \times m, s)$-nets and $(t, \alpha, \beta, \sigma, s)$-sequences can be constructed. The idea for the construction is based on the construction method presented in [2].

Let $d \ge 1$ and let $C_1, \ldots, C_{sd}$ be the generating matrices of a digital (t, m, sd)-net. Note that many explicit examples of such generating matrices are known, see for example [9, 20, 23, 31] and the references therein. For the construction of a (t, α, m, s)-net any of the above mentioned explicit constructions can be used, but as will be shown below the quality of the (t, α, m, s)-net obtained depends on the quality of the underlying digital (t, m, sd)-net on which our construction is based on.

Let $C_j = (c_{j,1}, \ldots, c_{j,m})^\top$ for $j = 1, \ldots, sd$, i.e., $c_{j,l}$ are the row vectors of C_j. Now let the matrix $C_j^{(d)}$ be made of the first rows of the matrices $C_{(j-1)d+1}, \ldots, C_{jd}$, then the second rows of $C_{(j-1)d+1}, \ldots, C_{jd}$ and so on. The matrix $C_j^{(d)}$ is then an $dm \times m$ matrix, i.e., $C_j^{(d)} = (c_{j,1}^{(d)}, \ldots, c_{j,dm}^{(d)})^\top$ where $c_{j,l}^{(d)} = c_{u,v}$ with $l = (v-j)d + u$, $1 \le v \le m$ and $(j-1)d < u \le jd$ for $l = 1, \ldots, dm$ and $j = 1, \ldots, s$. The following result is a slight generalization of [2, Theorem 3] and can be obtained using the same proof technique.

THEOREM 4.11. *Let $d \ge 1$ be a natural number and let $C_1, \ldots, C_{sd}$ be the generating matrices of a digital (t', m, sd)-net over the finite field $\mathbb{F}_q$ of prime power order q. Let $C_1^{(d)}, \ldots, C_s^{(d)}$ be defined as above. Then for any $\alpha \ge 1$ the matrices $C_1^{(d)}, \ldots, C_s^{(d)}$ are generating matrices of a digital $(t, \alpha, \min(1, \alpha/d), dm \times m, s)$-net over $\mathbb{F}_q$ with*

$$t = \min(\alpha, d)\, t' + \left\lceil \frac{s(d-1)\min(\alpha, d)}{2} \right\rceil.$$

The above construction and Theorem 4.11 can easily be extended to $(t, \alpha, \beta, \sigma, s)$-sequences. Indeed, let $d \geq 1$ and let $C_1, \ldots, C_{sd}$ be the generating matrices of a digital (t, sd)-sequence. Again many explicit generating matrices are known, see for example [9, 20, 23, 31]. Let $C_j = (c_{j,1}, c_{j,2}, \ldots)^\top$ for $j = 1, \ldots, sd$, i.e., $c_{j,l}$ are the row vectors of C_j. Now let the matrix $C_j^{(d)}$ be made of the first rows of the matrices $C_{(j-1)d+1}, \ldots, C_{jd}$, then the second rows of $C_{(j-1)d+1}, \ldots, C_{jd}$ and so on, i.e.,

$$C_j^{(d)} = (c_{(j-1)d+1,1}, \ldots, c_{jd,1}, c_{(j-1)d+1,2}, \ldots, c_{jd,2}, \ldots)^\top.$$

The following theorem states that the matrices $C_1^{(d)}, \ldots, C_s^{(d)}$ are the generating matrices of a digital $(t, \alpha, \min(1, \alpha/d), d, s)$-sequence, compare with [2, Theorem 4].

THEOREM 4.12. *Let $d \geq 1$ be a natural number and let $C_1, \ldots, C_{sd}$ be the generating matrices of a digital (t', sd)-sequence over the finite field $\mathbb{F}_q$ of prime power order q. Let $C_1^{(d)}, \ldots, C_s^{(d)}$ be defined as above. Then for any $\alpha \geq 1$ the matrices $C_1^{(d)}, \ldots, C_s^{(d)}$ are generating matrices of a digital $(t, \alpha, \min(1, \alpha/d), d, s)$-sequence over $\mathbb{F}_q$ with*

$$t = \min(\alpha, d) \, t' + \left\lceil \frac{s(d-1)\min(\alpha, d)}{2} \right\rceil.$$

The last result shows that $(t, \alpha, \beta, \sigma m \times m, s)$-nets indeed exist for $\beta = 1$ and any $0 < \sigma \leq \alpha$ and for m arbitrarily large. We have even shown that digital $(t, \alpha, \beta, \alpha m \times m, s)$-nets exist which are extensible in m and s. This can be achieved by using an underlying (t', sd)-sequence which is itself extensible in m and s. If the t' value of the original (t', m, s)-net or (t', s)-sequence is known explicitly then we also know the t value of the digital $(t, \alpha, \beta, \alpha m \times m, s)$-net or $(t, \alpha, \beta, \sigma, s)$-sequence. Furthermore it has also been shown how such digital nets can be constructed in practice. Further results on such sequences are established in [2].

5. Numerical integration in the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$. In this section we investigate numerical integration in the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ using quasi-Monte Carlo rules

$$Q_{q^m,s}(f) = \frac{1}{q^m} \sum_{n=0}^{q^m-1} f(\boldsymbol{x}_n),$$

where $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{q^m-1}$ are the points of a digital (t, α, β, m, s)-net over $\mathbb{F}_q$. More precisely, we want to approximate the integral

$$I_s(f) = \int_{[0,1]^s} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$$

by the quasi-Monte Carlo rule $Q_{q^m,s}(f)$. As a quality measure for our rule we introduce the worst-case error in the next section.

5.1. The worst-case error in the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$. The worst-case error for the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ using the quasi-Monte Carlo rule $Q_{q^m,s}$ is given by

$$e(Q_{q^m,s}, \mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}) = \sup_{\substack{f \in \mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}} \\ \|f\|_{\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}} \leq 1}} |I_s(f) - Q_{q^m,s}(f)|.$$

The initial error is given by

$$e(Q_{0,s}, \mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}) = \sup_{\substack{f \in \mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}} \\ \|f\|_{\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}} \leq 1}} |I_s(f)|.$$

In the following we use digital nets generated by the matrices $C_1, \ldots, C_s$ as quadrature points for the quadrature rule $Q_{q^m,s}$. Let $f \in \mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$. Using Lemma 4.2 it follows that

$$|I_s(f) - Q_{q^m,s}(f)| = \left| \sum_{\boldsymbol{k} \in \mathcal{D}} \hat{f}(\boldsymbol{k}) \right|$$

$$\leq \sum_{\boldsymbol{k} \in \mathcal{D}} |\hat{f}(\boldsymbol{k})| = \sum_{\emptyset \neq u \subseteq \mathcal{S}} \sum_{\boldsymbol{k}_u \in \mathcal{D}_u^*} |\hat{f}(\boldsymbol{k}_u, \boldsymbol{0}_{\mathcal{S} \setminus u})|.$$

Now we have $|\hat{f}(\boldsymbol{k}_u, \boldsymbol{0}_{\mathcal{S} \setminus u})| \leq \gamma_u r_{q,\vartheta}(\boldsymbol{k}_u) \|f\|_{\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}}$ and thus we obtain

$$|I_s(f) - Q_{q^m,s}(f)| \leq \|f\|_{\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}} \sum_{\emptyset \neq u \subseteq \mathcal{S}} \gamma_u \sum_{\boldsymbol{k}_u \in \mathcal{D}_u^*} r_{q,\vartheta}(\boldsymbol{k}_u). \tag{5.1}$$

By choosing $\hat{f}(\boldsymbol{k}_u, \boldsymbol{0}_{\mathcal{S} \setminus u}) = \gamma_u r_{q,\vartheta}(\boldsymbol{k}_u)$ for all u and $\boldsymbol{k}_u$ we can also obtain equality in (5.1). Thus we have

$$e(Q_{q^m,s}, \mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}) = \sum_{\emptyset \neq u \subseteq \mathcal{S}} \gamma_u \sum_{\boldsymbol{k}_u \in \mathcal{D}_u^*} r_{q,\vartheta}(\boldsymbol{k}_u). \tag{5.2}$$

From the last formula we can now see that essentially a large value of $\min\{\mu_{q,\vartheta}(\boldsymbol{k}) : \boldsymbol{k} \in \mathcal{D}\}$ guarantees a small worst-case error. Further it can be shown that

$$e(Q_{0,s}, \mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}) = \gamma_\emptyset. \tag{5.3}$$

We have shown the following theorem.

THEOREM 5.1. *The initial error for multivariate integration in the Walsh space* $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ *is given by (5.3) and the worst-case error for multivariate integration in the Walsh space* $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ *using a digital net as quadrature points is given by (5.2).*

In the following lemma we establish an upper bound on the sum $\sum_{\boldsymbol{k}_u \in \mathcal{D}_u^*} r_{q,\vartheta}(\boldsymbol{k}_u)$ for digital $(t, \alpha, \beta, n \times m, s)$-nets over $\mathbb{F}_q$. The proof is similar to [2, Lemma 6].

LEMMA 5.2. *Let* $\vartheta > 1$ *be a real number,* $q \geq 2$ *be a prime power,* $C_1, \ldots, C_s \in \mathbb{F}_q^{n \times m}$ *be the generating matrices of a digital* $(t, \lceil \vartheta \rceil, \beta, n \times m, s)$-net over $\mathbb{F}_q$ with $0 < \beta \leq 1$ and let $\mathcal{D}_u^* = \mathcal{D}_u^*((C_j)_{j \in u})$. For all $\emptyset \neq u \subseteq \mathcal{S}$ we have: if ϑ is not an integer it follows that

$$\sum_{\boldsymbol{k}_u \in \mathcal{D}_u^*} r_{q,\vartheta}(\boldsymbol{k}_u) \leq C_{|u|,q,\vartheta}(\beta n - t + \lceil \vartheta \rceil)^{|u| \lceil \vartheta \rceil - 1} q^{-\vartheta \lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor},$$

where

$$C_{|u|,q,\vartheta} = q^{|u| \lceil \vartheta \rceil} ((q - q^{\vartheta - \lfloor \vartheta \rfloor})^{-1} + (1 - q^{(1-\vartheta)/\lceil \vartheta \rceil})^{-|u| \lceil \vartheta \rceil})$$

and if ϑ *is an integer it follows that*

$$\sum_{\boldsymbol{k}_u \in \mathcal{D}_u^*} r_{q,\vartheta}(\boldsymbol{k}_u) \leq C'_{|u|,q,\vartheta}(\beta n - t + \vartheta)^{|u| \vartheta} q^{-(\beta n - t)},$$

where

$$C'_{|u|,q,\vartheta} = q^{|u|\vartheta}(q^{-1} + (1 - q^{1/\vartheta-1})^{-|u|\vartheta}).$$

Proof. To simplify the notation we prove the result only for $u = \mathcal{S}$. For all other subsets the result follows by the same arguments.

We first consider the case where $\vartheta > 1$ is not an integer. We partition the set $\mathcal{D}_\mathcal{S}^*$ into parts where the highest digits of k_j are prescribed and we count the number of solutions of $C_1^\top \vec{k}_1 + \cdots + C_s^\top \vec{k}_s = \vec{0}$. For $j = 1, \ldots, s$ let now $i_{j,\lceil\vartheta\rceil} < \cdots < i_{j,1}$ with $i_{j,1} \geq 1$. Note that we now allow $i_{j,l} < 1$, in which case the contributions of those $i_{j,l}$ are to be ignored. This notation is adopted in order to avoid considering many special cases. Further we write $\boldsymbol{i}_{s,\lceil\vartheta\rceil} = (i_{1,1}, \ldots, i_{1,\lceil\vartheta\rceil}, \ldots, i_{s,1}, \ldots, i_{s,\lceil\vartheta\rceil})$ and define

$$\mathcal{D}_\mathcal{S}^*(\boldsymbol{i}_{s,\lceil\vartheta\rceil}) = \{\boldsymbol{k} \in \mathcal{D}_\mathcal{S}^* : k_j = \lfloor \kappa_{j,1}q^{i_{j,1}-1} + \cdots + \kappa_{j,\lceil\vartheta\rceil}q^{i_{j,\lceil\vartheta\rceil}-1} + l_j \rfloor$$
$$\text{with } 0 \leq l_j < q^{i_{j,\lceil\vartheta\rceil}-1} \text{and } 1 \leq \kappa_{j,l} < q \text{ for } j = 1, \ldots, s\},$$

where $\lfloor \cdot \rfloor$ just means that the contributions of $i_{j,l} < 1$ are to be ignored. Let $\mu(\boldsymbol{i}_{s,\lceil\vartheta\rceil}) = i_{1,1} + \cdots + i_{1,\lceil\vartheta\rceil-1} + (\vartheta - \lfloor\vartheta\rfloor)i_{1,\lceil\vartheta\rceil} + \cdots + i_{s,1} + \cdots + i_{s,\lceil\vartheta\rceil-1} + (\vartheta - \lfloor\vartheta\rfloor)i_{s,\lceil\vartheta\rceil}$.
Then we have

$$\sum_{\boldsymbol{k}_\mathcal{S} \in \mathcal{D}_\mathcal{S}^*} r_{q,\vartheta}(\boldsymbol{k}_\mathcal{S}) = \sum_{i_{1,1}=1}^{\infty} \cdots \sum_{i_{1,\lceil\vartheta\rceil}=1}^{i_{1,\lceil\vartheta\rceil-1}-1} \cdots \sum_{i_{s,1}=1}^{\infty} \cdots \sum_{i_{s,\lceil\vartheta\rceil}=1}^{i_{s,\lceil\vartheta\rceil-1}-1} \frac{|\mathcal{D}_\mathcal{S}^*(\boldsymbol{i}_{s,\lceil\vartheta\rceil})|}{q^{\mu(\boldsymbol{i}_{s,\lceil\vartheta\rceil})}}. \quad (5.4)$$

Some of the sums above can be empty in which case we just set the corresponding summation index $i_{j,l} = 0$.

Note that by the $(t, \lceil\vartheta\rceil, \beta, n \times m, s)$-net property we have that $|\mathcal{D}_\mathcal{S}^*(\boldsymbol{i}_{s,\lceil\vartheta\rceil})| = 0$ as long as $i_{1,1} + \cdots + i_{1,\lceil\vartheta\rceil} + \cdots + i_{s,1} + \cdots + i_{s,\lceil\vartheta\rceil} \leq \beta n - t$. Hence let now $0 \leq i_{1,1}, \ldots, i_{s,\lceil\vartheta\rceil}$ be given such that $i_{1,1}, \ldots, i_{s,1} \geq 1$, $i_{j,\lceil\vartheta\rceil} < \cdots < i_{j,1}$ for $j = 1, \ldots, s$ and where if $i_{j,l} < 1$ we set $i_{j,l} = 0$ (in which case we also have $i_{j,l+1} = i_{j,l+2} = \ldots = 0$ and the inequalities $i_{j,l} > \cdots > i_{j,\lceil\vartheta\rceil}$ are ignored) and $i_{1,1} + \cdots + i_{1,\lceil\vartheta\rceil} + \cdots + i_{s,1} + \cdots + i_{s,\lceil\vartheta\rceil} > \beta n - t$. We now need to estimate $|\mathcal{D}_\mathcal{S}^*(\boldsymbol{i}_{s,\lceil\vartheta\rceil})|$, that is we need to count the number of $\boldsymbol{k} \in \mathcal{D}_\mathcal{S}^*$ with $k_j = \lfloor \kappa_{j,1}b^{i_{j,1}-1} + \cdots + \kappa_{j,\lceil\vartheta\rceil}b^{i_{j,\lceil\vartheta\rceil}-1} + l_j \rfloor$.

There are at most $(q-1)^{\lceil\vartheta\rceil s}$ choices for $\kappa_{1,1}, \ldots, \kappa_{s,\lceil\vartheta\rceil}$ (we write at most because if $i_{j,l} < 1$ then the corresponding $\kappa_{j,l}$ does not have any effect and therefore need not to be included).

Let now $1 \leq \kappa_{1,1}, \ldots, \kappa_{s,\lceil\vartheta\rceil} < q$ be given and define

$$\vec{g} = \kappa_{1,1}c_{1,i_{1,1}}^\top + \cdots + \kappa_{1,\lceil\vartheta\rceil}c_{1,i_{1,\lceil\vartheta\rceil}}^\top + \cdots + \kappa_{s,1}c_{s,i_{s,1}}^\top + \cdots + \kappa_{s,\lceil\vartheta\rceil}c_{s,i_{s,\lceil\vartheta\rceil}}^\top,$$

where we set $c_{j,l}^\top = 0$ if $l < 1$ or $l > n$. Further let

$$B = (c_{1,1}^\top, \ldots, c_{1,i_{1,\lceil\vartheta\rceil}-1}^\top, \ldots, c_{s,1}^\top, \ldots, c_{s,i_{s,\lceil\vartheta\rceil}-1}^\top).$$

Now the task is to count the number of solutions $\vec{l}$ of $B\vec{l} = \vec{g}$.

As long as the columns of B are linearly independent the number of solutions can at most be 1. By the $(t, \lceil\vartheta\rceil, \beta, n \times m, s)$-net property this is certainly the case if (we write $(x)_+ = \max(x, 0)$)

$$(i_{1,\lceil\vartheta\rceil} - 1)_+ + \cdots + (i_{1,\lceil\vartheta\rceil} - \lceil\vartheta\rceil)_+ + \cdots$$
$$+ (i_{s,\lceil\vartheta\rceil} - 1)_+ + \cdots + (i_{s,\lceil\vartheta\rceil} - \lceil\vartheta\rceil)_+ \leq \lceil\vartheta\rceil(i_{1,\lceil\vartheta\rceil} + \cdots + i_{s,\lceil\vartheta\rceil})$$
$$\leq \beta n - t,$$

that is, as long as

$$i_{1,\lceil\vartheta\rceil} + \cdots + i_{s,\lceil\vartheta\rceil} \le \frac{\beta n - t}{\lceil\vartheta\rceil}.$$

Let now $i_{1,\lceil\vartheta\rceil} + \cdots + i_{s,\lceil\vartheta\rceil} > \frac{\beta n-t}{\lceil\vartheta\rceil}$. Then by considering the rank of the matrix B and the dimension of the space of solutions of $B\vec{l} = \vec{0}$ it follows the number of solutions of $B\vec{l} = \vec{g}$ is smaller or equal to $q^{i_{1,\lceil\vartheta\rceil}+\cdots+i_{s,\lceil\vartheta\rceil}-\lfloor(\beta n-t)/\lceil\vartheta\rceil\rfloor}$. Thus we have

$$|\mathcal{D}_{\mathcal{S}}^*(\boldsymbol{i}_{s,\lceil\vartheta\rceil})| = 0$$

if $\sum_{j=1}^s \sum_{l=1}^{\lceil\vartheta\rceil} i_{j,l} \le \beta n - t$, we have

$$|\mathcal{D}_{\mathcal{S}}^*(\boldsymbol{i}_{s,\lceil\vartheta\rceil})| = (q-1)^{s\lceil\vartheta\rceil}$$

if $\sum_{j=1}^s \sum_{l=1}^{\lceil\vartheta\rceil} i_{j,l} > \beta n - t$ and $\sum_{j=1}^s i_{j,\lceil\vartheta\rceil} \le \frac{\beta n-t}{\lceil\vartheta\rceil}$ and finally we have

$$|\mathcal{D}_{\mathcal{S}}^*(\boldsymbol{i}_{s,\lceil\vartheta\rceil})| \le (q-1)^{s\lceil\vartheta\rceil} q^{i_{1,\lceil\vartheta\rceil}+\cdots+i_{s,\lceil\vartheta\rceil}-\lfloor(\beta n-t)/\lceil\vartheta\rceil\rfloor}$$

if $\sum_{j=1}^s \sum_{l=1}^{\lceil\vartheta\rceil} i_{j,l} > \beta n - t$ and $\sum_{j=1}^s i_{j,\lceil\vartheta\rceil} > \frac{\beta n-t}{\lceil\vartheta\rceil}$.

We estimate the sum (5.4) now. Let S_1 be the sum in (5.4) where $i_{1,1}+\cdots+i_{s,\lceil\vartheta\rceil} > \beta n - t$ and $i_{1,\lceil\vartheta\rceil} + \cdots + i_{s,\lceil\vartheta\rceil} \le \frac{\beta n-t}{\lceil\vartheta\rceil}$. Let $l_1 = i_{1,1} + \cdots + i_{1,\lceil\vartheta\rceil-1} + \cdots + i_{s,1} + \cdots + i_{s,\lceil\vartheta\rceil-1}$ and let $l_2 = i_{1,\lceil\vartheta\rceil} + \cdots + i_{s,\lceil\vartheta\rceil}$. Let $A(l_1 + l_2)$ denote the number of admissible choices of $i_{1,1}, \ldots, i_{s,\lceil\vartheta\rceil}$ such that $l_1+l_2 = i_{1,1}+\cdots+i_{s,\lceil\vartheta\rceil}$. Then we have

$$S_1 = (q-1)^{s\lceil\vartheta\rceil} \sum_{l_2=0}^{\lfloor\frac{\beta n-t}{\lceil\vartheta\rceil}\rfloor} \frac{1}{q^{(\vartheta-\lfloor\vartheta\rfloor)l_2}} \sum_{l_1=\beta n-t+1-l_2}^{\infty} \frac{A(l_1+l_2)}{b^{l_1}}.$$

We have $A(l_1 + l_2) \le \binom{l_1+l_2+s\lceil\vartheta\rceil-1}{s\lceil\vartheta\rceil-1}$ and hence we obtain

$$S_1 \le (q-1)^{s\lceil\vartheta\rceil} \sum_{l_2=0}^{\lfloor\frac{\beta n-t}{\lceil\vartheta\rceil}\rfloor} \frac{1}{q^{(\vartheta-\lfloor\vartheta\rfloor)l_2}} \sum_{l_1=\beta n-t+1-l_2}^{\infty} \frac{1}{q^{l_1}} \binom{l_1 + l_2 + s\lceil\vartheta\rceil - 1}{s\lceil\vartheta\rceil - 1}.$$

From a result by Matoušek [17, Lemma 2.18], see also [4, Lemma 6], we have

$$(q-1)^{s\lceil\vartheta\rceil} \sum_{l_1=\beta n-t+1-l_2}^{\infty} \frac{1}{q^{l_1}} \binom{l_1 + l_2 + s\lceil\vartheta\rceil - 1}{s\lceil\vartheta\rceil - 1}$$
$$\le q^{l_2-\beta n+t-1+s\lceil\vartheta\rceil} \binom{\beta n - t + s\lceil\vartheta\rceil}{s\lceil\vartheta\rceil - 1}$$

and further we have

$$\sum_{l_2=0}^{\lfloor\frac{\beta n-t}{\lceil\vartheta\rceil}\rfloor} \frac{q^{l_2}}{q^{(\vartheta-\lfloor\vartheta\rfloor)l_2}} = \sum_{l_2=0}^{\lfloor\frac{\beta n-t}{\lceil\vartheta\rceil}\rfloor} q^{l_2(\lceil\vartheta\rceil-\vartheta)} = \frac{q^{(\lceil\vartheta\rceil-\vartheta)(\lfloor(\beta n-t)/\lceil\vartheta\rceil\rfloor+1)} - 1}{q^{\lceil\vartheta\rceil-\vartheta} - 1}.$$

Thus we obtain

$$S_1 \leq \frac{q^{(\lceil \vartheta \rceil - \vartheta)(\lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor + 1)} - 1}{q^{\lceil \vartheta \rceil - \vartheta} - 1} q^{-\beta n + t - 1 + s\lceil \vartheta \rceil} \binom{\beta n - t + s\lceil \vartheta \rceil}{s\lceil \vartheta \rceil - 1}$$

$$\leq \frac{q^{s\lceil \vartheta \rceil - 1}}{1 - q^{\vartheta - \lceil \vartheta \rceil}} \binom{\beta n - t + s\lceil \vartheta \rceil}{s\lceil \vartheta \rceil - 1} q^{-\vartheta \lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor}.$$

Let S_2 be the part of (5.4) for which $i_{1,1} + \cdots + i_{s,\lceil \vartheta \rceil} > \beta n - t$ and $i_{1,\lceil \vartheta \rceil} + \cdots + i_{s,\lceil \vartheta \rceil} > \frac{\beta n - t}{\lceil \vartheta \rceil}$, i.e., we have

$$S_2 \leq (q-1)^{s\lceil \vartheta \rceil} \sum_{i_{1,1}=1}^{\infty} \cdots \sum_{i_{1,\lceil \vartheta \rceil}=1}^{i_{1,\lceil \vartheta \rceil - 1} - 1} \cdots$$

$$\sum_{i_{s,1}=1}^{\infty} \cdots \sum_{i_{s,\lceil \vartheta \rceil}=1}^{i_{s,\lceil \vartheta \rceil - 1} - 1} \frac{q^{-\lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor} q^{(i_{1,\lceil \vartheta \rceil} + \cdots + i_{s,\lceil \vartheta \rceil})(\lceil \vartheta \rceil - \vartheta)}}{q^{i_{1,1} + \cdots + i_{1,\lceil \vartheta \rceil - 1} + \cdots + i_{s,1} + \cdots + i_{s,\lceil \vartheta \rceil - 1}}},$$

where we have the additional conditions $i_{1,1} + \cdots + i_{s,\lceil \vartheta \rceil} > \beta n - t$ and $i_{1,\lceil \vartheta \rceil} + \cdots + i_{s,\lceil \vartheta \rceil} > \frac{\beta n - t}{\lceil \vartheta \rceil}$. As above let $l_1 = i_{1,1} + \cdots + i_{1,\lceil \vartheta \rceil - 1} + \cdots + i_{s,1} + \cdots + i_{s,\lceil \vartheta \rceil - 1}$ and let $l_2 = i_{1,\lceil \vartheta \rceil} + \cdots + i_{s,\lceil \vartheta \rceil}$. Let $A(l_1 + l_2)$ denote the number of admissible choices of $i_{1,1}, \ldots, i_{s,\lceil \vartheta \rceil}$ such that $l_1 + l_2 = i_{1,1} + \cdots + i_{s,\lceil \vartheta \rceil}$. Note that $l_1 > \lfloor \vartheta \rfloor l_2$. Then we have $A(l_1 + l_2) \leq \binom{l_1 + l_2 + s\lceil \vartheta \rceil - 1}{s\lceil \vartheta \rceil - 1}$ and hence we obtain

$$S_2 \leq (q-1)^{s\lceil \vartheta \rceil} q^{-\lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor}$$

$$\sum_{l_2 = \lfloor \frac{\beta n - t}{\lceil \vartheta \rceil} \rfloor + 1}^{\infty} q^{(\lceil \vartheta \rceil - \vartheta) l_2} \sum_{l_1 = \lfloor \vartheta \rfloor l_2 + 1}^{\infty} \frac{1}{q^{l_1}} \binom{l_1 + l_2 + s\lceil \vartheta \rceil - 1}{s\lceil \vartheta \rceil - 1}$$

$$= (q-1)^{s\lceil \vartheta \rceil} q^{-\lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor}$$

$$\sum_{l_2 = \lfloor \frac{\beta n - t}{\lceil \vartheta \rceil} \rfloor + 1}^{\infty} \sum_{l_1 = 0}^{\infty} q^{-l_1 + l_2 - 1 - l_2 \vartheta} \binom{l_1 + l_2 + \lfloor \vartheta \rfloor l_2 - 1 + s\lceil \vartheta \rceil - 1}{s\lceil \vartheta \rceil - 1}.$$

By using again Matoušek [17, Lemma 2.18], see also [4, Lemma 6], we have

$$(q-1)^{s\lceil \vartheta \rceil} \sum_{l_1 = 0}^{\infty} q^{-l_1 + l_2 - 1 - l_2 \vartheta} \binom{l_1 + l_2 + \lfloor \vartheta \rfloor l_2 - 1 + s\lceil \vartheta \rceil - 1}{s\lceil \vartheta \rceil - 1}$$

$$\leq q^{s\lceil \vartheta \rceil} q^{l_2(1 - \vartheta) - 1} \binom{l_2 \lceil \vartheta \rceil - 1 + s\lceil \vartheta \rceil - 1}{s\lceil \vartheta \rceil - 1}$$

and also

$$q^{s\lceil \vartheta \rceil - 1 - \lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor} \sum_{l_2 = \lfloor \frac{\beta n - t}{\lceil \vartheta \rceil} \rfloor + 1}^{\infty} q^{l_2(1 - \vartheta)} \binom{l_2 \lceil \vartheta \rceil - 1 + s\lceil \vartheta \rceil - 1}{s\lceil \vartheta \rceil - 1}$$

$$\leq q^{s\lceil \vartheta \rceil - 1 - \lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor} \sum_{l_2 = \beta n - t}^{\infty} q^{l_2(1 - \vartheta)/\lceil \vartheta \rceil} \binom{l_2 + \lceil \vartheta \rceil - 1 + s\lceil \vartheta \rceil - 1}{s\lceil \vartheta \rceil - 1}$$

$$\leq q^{s\lceil \vartheta \rceil} (1 - q^{(1 - \vartheta)/\lceil \vartheta \rceil})^{-s\lceil \vartheta \rceil} \binom{\beta n - t + \lceil \vartheta \rceil - 2 + s\lceil \vartheta \rceil}{s\lceil \vartheta \rceil - 1} q^{-\vartheta(\beta n - t)/\lceil \vartheta \rceil}.$$

Hence we have

$$S_2 \leq q^{s\lceil\vartheta\rceil}(1-q^{(1-\vartheta)/\lceil\vartheta\rceil})^{-s\lceil\vartheta\rceil}\binom{\beta n - t + \lceil\vartheta\rceil - 2 + s\lceil\vartheta\rceil}{s\lceil\vartheta\rceil - 1}q^{-\vartheta(\beta n - t)/\lceil\vartheta\rceil}.$$

Note that we have $\sum_{\boldsymbol{k}_S \in \mathcal{D}_S^*} r_{q,\vartheta}(\boldsymbol{k}_S) = S_1 + S_2$. Let $a \geq 1$ and $b \geq 0$ be integers, then we have

$$\binom{a+b}{b} = \prod_{i=1}^{b}\left(1 + \frac{a}{i}\right) \leq (1+a)^b.$$

Therefore we obtain

$$S_1 \leq \frac{q^{s\lceil\vartheta\rceil - 1}}{1 - q^{\vartheta - \lceil\vartheta\rceil}}(\beta n - t + 2)^{s\lceil\vartheta\rceil - 1}q^{-\vartheta\lfloor(\beta n - t)/\lceil\vartheta\rceil\rfloor}$$

and

$$S_2 \leq q^{s\lceil\vartheta\rceil}(1-q^{(1-\vartheta)/\lceil\vartheta\rceil})^{-s\lceil\vartheta\rceil}(\beta n - t + \lceil\vartheta\rceil)^{s\lceil\vartheta\rceil - 1}q^{-\vartheta(\beta n - t)/\lceil\vartheta\rceil}.$$

Thus we have

$$\sum_{\boldsymbol{k}_S \in \mathcal{D}_S^*} r_{q,\vartheta}(\boldsymbol{k}_S) \leq C_{s,q,\vartheta}(\beta n - t + \lceil\vartheta\rceil)^{s\lceil\vartheta\rceil - 1}q^{-\vartheta\lfloor(\beta n - t)/\lceil\vartheta\rceil\rfloor},$$

where

$$C_{s,q,\vartheta} = q^{s\lceil\vartheta\rceil}((q - q^{\vartheta - \lfloor\vartheta\rfloor})^{-1} + (1 - q^{(1-\vartheta)/\lceil\vartheta\rceil})^{-s\lceil\vartheta\rceil}).$$

The result follows for the case $0 < \vartheta - \lfloor\vartheta\rfloor < 1$.

Let now $\vartheta > 1$ be an integer. Then using the same arguments as above it can be shown that

$$S_1 \leq (\beta n - t + 2)^{s\vartheta}q^{-(\beta n - t) - 1 + s\vartheta}$$

and

$$S_2 \leq q^{s\vartheta}(1 - q^{1/\vartheta - 1})^{-s\vartheta}(\beta n - t + \vartheta)^{s\vartheta - 1}q^{-(\beta n - t)}.$$

Thus we have

$$\sum_{\boldsymbol{k}_S \in \mathcal{D}_S^*} r_{q,\vartheta}(\boldsymbol{k}_S) \leq C'_{s,q,\vartheta}(\beta n - t + \vartheta)^{s\vartheta}q^{-(\beta n - t)},$$

where

$$C'_{s,q,\vartheta} = q^{s\vartheta}(q^{-1} + (1 - q^{1/\vartheta - 1})^{-s\vartheta}).$$

The result now follows. $\square$

REMARK 5.3. We note that the above lemma does not hold for $\beta > 1$ in general. Indeed, take for example $u = \{1\}$, then $\boldsymbol{k}_u = (k_1)$ and choose $k_1 = q^n$. Then the digit vector of the first n digits of q^n is $(0, \ldots, 0)^\top$ and hence $C_1^\top \vec{k}_1 = \vec{0}$ and hence $\boldsymbol{k}_{(1)} \in \mathcal{D}_{(1)}^*$. Thus

$$\sum_{\boldsymbol{k}_{(1)} \in \mathcal{D}_{(1)}^*} r_{q,\vartheta}(\boldsymbol{k}_{(1)}) \geq q^{-n-1}$$

and hence a counterexample can be obtained for some choices of n, β, ϑ.

In [2] we did allow $\beta > 1$, but therein we had the additional assumption that the functions are periodic. In this case we were able to show that the Walsh coefficients $r_{q,\alpha}(\boldsymbol{k}, \boldsymbol{l}) = \prod_{j=1}^{s} r_{q,\alpha}(k_j, l_j)$ of the reproducing kernel also satisfy the additional property that $r_{q,\alpha}(q^m k_j, q^m k_j) = r_{q,\alpha}(q^m k_j) = q^{-2\alpha m} r_{q,\alpha}(k_j, k_j)$ for all $k_j, m \in \mathbb{N}$, see [2, Lemma 15]. Similarly, if we would also assume here that $r_{q,\vartheta}(q^n k) = q^{-\vartheta n} r_{q,\vartheta}(k)$ and $r_{q,\vartheta}(k)$ given as above if $q \nmid k$, then the above counterexample would fail as then $r_{q,\vartheta}(q^n) = r_{q,\vartheta}(1q^n) = q^{-\vartheta(n+1)} r_{q,\vartheta}(1)$.

Using the above lemma we can now obtain an upper bound on the worst-case error.

THEOREM 5.4. *Let $\vartheta > 1$ be a real number and $q \geq 2$ be a prime power. The worst-case error for multivariate integration in the Walsh space $\mathcal{E}_{s,q,\vartheta,\gamma}$ using a digital $(t, \lceil \vartheta \rceil, \beta, n \times m, s)$-net over $\mathbb{F}_q$, with $0 < \beta \leq 1$, as quadrature points is for non-integers ϑ bounded by*

$$e(Q_{q^m,s}, \mathcal{E}_{s,q,\vartheta,\gamma}) \leq q^{-\vartheta \lfloor (\beta n - t)/\lceil \vartheta \rceil \rfloor} \sum_{\emptyset \neq u \subseteq \mathcal{S}} \gamma_u C_{|u|,q,\vartheta} (\beta n - t + \lceil \vartheta \rceil)^{|u| \lceil \vartheta \rceil - 1},$$

where

$$C_{|u|,q,\vartheta} = q^{|u| \lceil \vartheta \rceil} ((q - q^{\vartheta - \lfloor \vartheta \rfloor})^{-1} + (1 - q^{(1-\vartheta)/\lceil \vartheta \rceil})^{-|u| \lceil \vartheta \rceil}),$$

and if ϑ is an integer, the worst-case error is bounded by

$$e(Q_{q^m,s}, \mathcal{E}_{s,q,\vartheta,\gamma}) \leq q^{-(\beta n - t)} \sum_{\emptyset \neq u \subseteq \mathcal{S}} \gamma_u C'_{|u|,q,\vartheta} (\beta n - t + \vartheta)^{|u| \vartheta},$$

where

$$C'_{|u|,q,\vartheta} = q^{|u| \vartheta} (q^{-1} + (1 - q^{1/\vartheta - 1})^{-|u| \vartheta}).$$

As a direct consequence of Corollary 3.11 we obtain the following result.

COROLLARY 5.5. *Let $\delta \geq 1$ be an integer, $0 < \lambda \leq 1$ and $q \geq 2$ be a prime power. Then for any function $f : [0,1)^s \to \mathbb{R}$ whose partial mixed derivatives up to order δ exist it follows that the integration error using a digital $(t, \delta + 1, \beta, n \times m, s)$-net over $\mathbb{F}_q$ with $0 < \beta \leq 1$ as quadrature points is for $0 < \lambda < 1$ bounded by*

$$|I_s(f) - Q_{q^m,s}(f)| \leq q^{-(\delta + \lambda) \lfloor (\beta n - t)/(\delta + 1) \rfloor} C_{\delta,s,q,\gamma} N_{\delta,\lambda,\gamma}(f)$$
$$\sum_{\emptyset \neq u \subseteq \mathcal{S}} \gamma_u C_{|u|,q,\delta+\lambda} (\beta n - t + \delta + 1)^{|u|(\delta+1)-1}$$

and for $\lambda = 1$ the integration error is bounded by

$$|I_s(f) - Q_{q^m,s}(f)| \leq q^{-(\beta n - t)} C_{\delta,s,q,\gamma} N_{\delta,\lambda,\gamma}(f)$$
$$\sum_{\emptyset \neq u \subseteq \mathcal{S}} \gamma_u C'_{|u|,q,\delta+1} (\beta n - t + \delta + 1)^{|u|\delta+1},$$

where the constant $C_{\delta,s,q,\gamma}$ is given in Corollary 3.11 and the constants $C_{|u|,q,\delta+\lambda}$ and $C'_{|u|,q,\delta+1}$ are given in Theorem 5.4.

Explicit constructions of digital $(t, \alpha, \min(1, \alpha/d), dm \times m, s)$-nets over $\mathbb{F}_q$ for all prime powers q, integers $\alpha, d, m, s > 1$ are given in Section 4.4. By choosing

$d = \alpha = \lceil \vartheta \rceil = \delta + 1$, by Theorem 5.4 and Corollary 5.5 we obtain a convergence of $\mathcal{O}(q^{-\vartheta m} m^{s\lceil \vartheta \rceil + 1})$, which is optimal even for the smooth functions contained in the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$, see [29] where a lower bound for smooth periodic functions was shown.

REMARK 5.6. In [2, Remark 4] it was noted that if $m = n$ and $\beta > \alpha$ the t-value must grow with m and hence the restriction $\beta \leq \alpha$ was added. A similar argument yields in our case that the t-value must grow with n if $\beta n > \alpha m$ as Theorem 5.4 shows a convergence of $\mathcal{O}(q^{-\beta n + t})$ but the best possible convergence rate is $q^{-\alpha m}$, hence the restriction $\beta \leq \alpha m / n$ was added.

In case the smoothness of the function is not known our constructions adjust themselves automatically up to a certain degree in the following way: for the construction of the digital net we choose some value of $d \geq 1$ and construct a digital $(t, \alpha, \min(1, \alpha/d), dm \times m, s)$-net or a digital $(t, \alpha, \min(1, \alpha/d), d, s)$-sequence for all $\alpha \geq 1$. The values $\delta \geq 1$ and $0 < \lambda \leq 1$ determine the real smoothness of the function, which we now assume is not known. The value of α is the smoothness analog for the digital net, i.e., we need to choose $\alpha = \delta + 1$. First assume that $\delta + \lambda \leq d$, then $\min(1, \alpha/d) = (\delta + 1)/d$ and therefore we have $\beta = (\delta + 1)/d$. As $n = dm$ it follows that $\beta n = (\delta + 1)m$ and therefore Corollary 5.5 shows that we achieve a convergence of $\mathcal{O}(q^{-(\delta+\lambda)m} m^{s(\delta+1)+1})$, which is optimal. Now assume on the other hand that $\delta + \lambda > d$, then $\min(1, \alpha/d) = 1$ and therefore $\beta = 1$. Again we have $n = dm$ and hence $\beta n = dm$. In this case Corollary 5.5 shows that our construction achieves a convergence of $\mathcal{O}(q^{-dm} m^{s(\delta+1)+1})$.

Note that numerical integration of functions with less smoothness, i.e., for example functions with partial mixed derivatives up to degree 1 in $\mathcal{L}_2([0,1)^s)$ or functions with bounded variation, has been considered in many papers and monographs, see for example [3, 4, 6, 7, 14, 20, 30, 31]. Using the notation from above, basically those results are concerned with the case where $\delta = 0$ and $\lambda = 1$, hence the results here are a direct continuation of what was previously known. The construction of digital nets proposed here for $d = 1$ yields obviously digital (t, m, s)-nets and (t, s)-sequences as for example defined in [20]. In view of Corollary 5.5 and the explanation which followed it is hence not surprising that the classical examples and theory (see for example [3, 4, 12, 13, 14, 20, 26, 31]) only yielded a convergence of $\mathcal{O}(q^{m(-1+\varepsilon)})$ for any $\varepsilon > 0$ (the ε here is used to hide the powers of m).

Note that the worst-case error in the Walsh space $\mathcal{E}_{s,q,\vartheta,\boldsymbol{\gamma}}$ is invariant with respect to a digital shift (see [4]), hence Corollary 5.5 also holds for digitally shifted digital nets. Thus, if one wants to use randomized digital nets, one can also use randomly digitally shifted digital nets. The root mean square worst-case error for this case would of course be bounded by the bound in Corollary 5.5, as this bound holds for any digital shift, i.e., our result here is even stronger in that we have shown that even for the worst digital shift we still have the bound of Corollary 5.5. From this, it follows that for our situation here, there is, in some sense, no bad digital shift. Other more sophisticated scrambling methods which do not destroy the essential properties of the point set can be used as well (for example a digital shift of depth m, see [5, 17]), see [24] for some ideas in this direction.

REFERENCES

[1] H.E. Chrestenson, A class of generalized Walsh functions, Pacific J. Math., 5 (1955), 17–31.
[2] J. Dick, Explicit constructions of quasi-Monte Carlo rules for the numerical integration of high dimensional periodic functions, SIAM J. Numer. Anal., 45(2007), 2141–2176.

[3] J. Dick, F.Y. Kuo, F. Pillichshammer and I.H. Sloan, Construction algorithms for polynomial lattice rules for multivariate integration, Math. Comp., 74 (2005), 1895–1921.

[4] J. Dick and F. Pillichshammer, Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces, J. Complexity, 21 (2005), 149–195.

[5] J. Dick and F. Pillichshammer, On the mean square weighted L_2 discrepancy of randomized digital (t, m, s)-nets over $\mathbb{Z}_2$, Acta Arith., 117 (2005), 371–403.

[6] J. Dick, I.H. Sloan, X. Wang and H. Woźniakowksi, Liberating the weights, J. Complexity, 20 (2004), 593–623.

[7] J. Dick, I.H. Sloan, X. Wang and H. Woźniakowski, Good lattice rules in weighted Korobov spaces with general weights, Numer. Math., 103 (2006), 63–97.

[8] B. Efron and C. Stein, The jackknife estimate of variance, Ann. Stat., 9 (1981), 586–598.

[9] H. Faure, Discrèpances de suites associèes á un système de numèration (en dimension s), Acta Arith., 41 (1982), 337–351.

[10] N.J. Fine, On the Walsh functions, Trans. Amer. Math. Soc., 65 (1949), 372–414.

[11] S. Heinrich, F.J. Hickernell and R.X. Yue, Optimal quadrature for Haar wavelet spaces, Math. Comp., 73 (2004), 259–277.

[12] E. Hlawka, Funktionen von beschränkter Variation in der Theorie der Gleichverteilung, Ann. Mat. Pura Appl., 54 (1961), 325–333.

[13] J.F. Koksma, Een algemeene stelling uit de theorie der gelijkmatige verdeeling modulo 1, Mathematica B (Zutphen), 11 (1942/43), 7–11.

[14] L. Kuipers and H. Niederreiter, Uniform distribution of sequences, Pure and Applied Mathematics, John Wiley, New York, 1974.

[15] G. Larcher, H. Niederreiter and W.Ch. Schmid, Digital nets and sequences constructed over finite rings and their application to quasi-Monte Carlo integration, Monatsh. Math., 121 (1996), 231–253.

[16] G. Larcher, G. Pirsic, Base change problems for generalized Walsh series and multivariate numerical integration, Pacific J. Math., 189 (1999), 75–105.

[17] J. Matoušek, Geometric Discrepancy, Algorithms and Combinatorics 18, Springer, Berlin, 1999.

[18] H. Niederreiter, Quasi-Monte Carlo methods and pseudo-random numbers, Bull. Amer. Math. Soc. 84 (1978), 957-1041.

[19] H. Niederreiter, Low-discrepancy point sets, Monatsh. Math., 102 (1986), 155–167.

[20] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, CBMS–NSF Series in Applied Mathematics 63, SIAM, Philadelphia, 1992.

[21] H. Niederreiter, Constructions of (t, m, s)-nets and (t, s)-sequences, Finite Fields Appl., 11 (2005), 578–600.

[22] H. Niederreiter and G. Pirsic, Duality for digital nets and its applications, Acta Arith., 97 (2001), 173–182.

[23] H. Niederreiter and C.P. Xing, Quasirandom points and global function fields, in Finite Fields and Applications, S. Cohen and H. Niederreiter, eds., London Math. Soc. Lecture Note Series, Vol. 233, Cambridge University Press, Cambridge, 1996, 269–296.

[24] A.B. Owen, Monte Carlo, quasi-Monte Carlo, and randomized quasi-Monte Carlo, in: Monte Carlo and quasi-Monte Carlo methods 1998 (Claremont, CA), Springer, Berlin, 2000, 86–97.

[25] G. Pirsic, Embedding theorems and numerical integration of Walsh series over groups, PhD thesis, University of Salzburg, 1997.

[26] G. Pirsic, J. Dick and F. Pillichshammer, Cyclic digital nets, hyperplane nets and multivariate integration in Sobolev spaces, SIAM J. Numer. Anal., 44 (2006), 385–411.

[27] H.A. Rademacher, Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen, Math. Ann., 87 (1922), 112–138.

[28] M.Yu. Rosenbloom and M.A. Tsfasman, Codes in the m-metric, Problemi Peredachi Inf., 33 (1997), 45–52.

[29] I.F. Sharygin, A lower estimate for the error of quadrature formulas for certain classes of functions, Zh. Vychisl. Mat. i Mat. Fiz., 3 (1963), 370–376.

[30] I.H. Sloan and H. Woźniakowski, When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?, J. Complexity, 14 (1998), 1–33.

[31] I.M. Sobol, The distribution of points in a cube and the approximate evaluation of integrals, Zh. Vychisl. Mat. i Mat. Fiz., 7 (1967), 784–802.

[32] J. Stoer and R. Bulirsch, Introduction to Numerical Analysis, Texts in Applied Mathematics 12, Springer, New York, 2002.

[33] J.L. Walsh, A closed set of normal orthogonal functions, Amer. J. Math., 55 (1923), 5–24.

[34] A. Zygmund, Trigonometric series, Cambridge University Press, Cambridge, 1959.